Iterative Algorithms for Volumetric X-ray Computed Tomography

Qiu, Wei

Award date:
2012

Awarding institution:
University of Bath

Iterative algorithms for volumetric X-ray computed tomography

Wei Qiu

Department of Electronic & Electrical Engineering

University of Bath

A thesis submitted for the degree of

Doctor of Philosophy

Bath, 2012

Copyright

I would like to dedicate this thesis to my loving parents ...

Acknowledgements

Firstly, I acknowledge the financial support from the University of Bath and my parents. More importantly, I would like to express my sincere gratitude to my supervisors, Dr. Manuchehr Soleimani and Prof. Cathryn N. Mitchell, for their selflessness support and encouragement. As my lead supervisor, Dr. Manuchehr Soleimani gave me countless help and professional suggestion from the very beginning. I greatly appreciate his selfless giving and the time spent on me and the project. And I would also like to thank my supervisor, Prof. Cathryn N. Mitchell, who gave me lots of encouragement, guidance and support when I was doing my project. I owe the success of the project to them. I am also thankful to Jenna Tong, Paul Spencer, Thanyawee Pengpan and other members in the Invert group, for their inspired discussion as the project progressed. Besides, I would like to thank Tom Marchant and Chris Moore from the Christie Hospital, who provided us measured data for experiments.

I have no words to express gratitude to my family, especially my parents. Without their support and encouragement, I would never have been able to study in United Kingdom and have finished this thesis. I owe what I have, entirely, and I would have in the future, to them - my beloved parents.

Abstract

Cone beam computed tomography (CBCT) enables a volumetric image reconstruction from a set of 2D projection data. This thesis studies the performance of a wide range iterative algorithms in various aspects, aiming to generate a better CBCT image reconstruction, especially when projection data is limited. We have implemented a wide range of algebraic iterative algorithms. Hence, the performance of ART, SART and OS-SART is studied based on a range of image quality measures. The major limitations of traditional iterative methods are their computational time. The conjugate gradients (CG) algorithm and its variants can be used to solve linear systems of equations arising from CBCT. Their applications can be found in a general linear algebra context, but in tomography problems (e.g. CBCT reconstruction) they have not widely been used. Hence, CBCT reconstruction using the CG-type algorithm LSQR was implemented and studied. In CBCT reconstruction, the main computational challenge is that the matrix A usually is very large, and storing it in full requires an amount of memory well beyond the reach of commodity computers. Because of these memory capacity constraints, only a small fraction of the weighting matrix A is typically used, leading to a poor reconstruction. In this final part of the thesis, to overcome this difficulty, the matrix A is partitioned and stored blockwise, and blockwise matrix-vector multiplications are implemented within LSQR. This implementation allows us to use the full weighting matrix A for CBCT reconstruction without further enhancing computer standards. Tikhonov regularization has been developed in this framework, and can produce significant improvement in the reconstructed images for limited data case.

Contents

List of Figures

Chapter 1

Introduction

In 1972, computed tomography (CT) scanner was invented. CT is the cross-sectional imaging of an object from data that is collected by scanning the object from source to detectors by many different directions. Thanks to the invention of G.N. Hounsfield, for this technique's revolutionary approach in diagnostic medicine, it is possible, for the first time, to enable doctors to inspect a person's inside anatomy without the need for invasive surgery. The basic idea of CT is that, with a very large number of measurements from different directions and angles, it is possible to reconstruct a cross-sectional slice of a patient. X-rays were then firstly utilised in medical imaging application based on their x-ray attenuation coefficient in different tissue materials. Later, radioisotopes, ultrasound and magnetic resonance were also applied to medical imaging, but the fun-

damental theories differ according to their unique natural properties.

The problem about how to reconstruct an image from its projections dates back to 1917 by Radon in [3]. In the last decades, the reconstructed techniques in the field was moving rapidly while there are two main reconstruction techniques in CT nowadays, which are analytical and algebraic approaches. Analytical approach is based on the Fourier Slice Theorem [4] while the other approach is iterative methods that solve the reconstruction problem by solving a system of simultaneous linear equations, which was used by Hounsfield for his x-ray scanner. For the analytical approach, the most prominent one is the Filtered Backprojection (FBP) algorithm firstly proposed by [5, 6] in 1967 and it was popularised later on by Shepp and Logan in [7]. In this algorithm, the image is reconstructed in the following steps. The measurement projection data in the frequency domain are filtered and then the filtered version of the data are backprojected to form the reconstructed image. Because the computational time cost of these algorithms is low, FBP like algorithms based on Feldkamp, Davis and Kress (FDK) algorithm [8] are now routinely available in commercial radiology and radiotherapy equipment. The representative of another approach, the algebraic type algorithms, is the Algebraic Reconstruction Technique (ART) [9]. For this algorithm, the image is

reconstructed by applying a recursive updating method, using the differences between projection and re-projection data to correct the image. In this process, the re-projection data is obtained by projecting the current state of the reconstructed image and then measure how close the calculated data match to the measurement projection data. After that, in the backprojection step a corrective value is implemented to generate a new reconstructed image. These steps are recursive used until an optimal point where, in an ideal case, the reconstructed image matches exactly the same as the original image.

In real life application, for example, CBCT is widely used in image guided radiation therapy (IGRT). A device rotates around patients' bodies to acquire updated volumetric data without requiring any movement of the patient. This is clearly important, as any movement due to angle or position changes of the patients' body may lead to healthy cells being exposed to radiation during the treatment process. Compared with high resolution high dose fan beam CT scans, a CBCT scan is relatively low resolution but also low dose. CBCT machine and the scanning process are shown in Figure 1.1.

To use less dose is the main concern in CBCT for IGRT especially this is implemented daily. This becomes the main motivation and aim for this thesis. In this thesis,

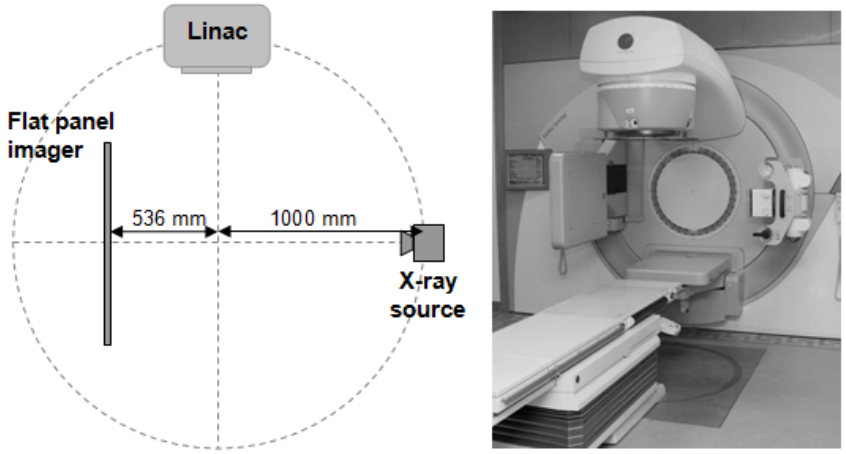


Figure 1.1: Medical linear accelerator with integrated X-ray cone-beam CT system.

works have been done in order to improve the speed of the CBCT reconstructions while reducing the amount of projection data used. These results will be shown in the following chapters.

The rest of the thesis is organised as the following. Chapter 2 will introduce the background related to CT reconstruction and some prior work will also be presented. In Chapter 3, various iterative methods used throughout the project will be presented. Image quality assessment will then be presented in Chapter 4. Parameter optimisation for multi-instrument data analysis system (MIDAS) based iterative methods is discussed in Chapter 5. In Chapter 6, convergence and image quality studies will be shown. Chapter 7 will introduce a modified iterative algorithm that can be used for

volumetric CT reconstruction. Finally, conclusions and future work will be made in Chapter 8.

Chapter 2

CBCT image reconstructions

Cone beam computed tomography (CBCT) is widely used in many real life applications, as it provides volumetric and high quality image reconstruction from large scale projection data. The fundamental theory of tomographic image is to reconstruct an image from its measurements or, more precisely, projections. A projection is the integral of the image in a direction specified by a given angle as illustrated in Figure 2.1. In this figure, objects are scanned by X-rays. In the detectors, projections are obtained. In the simplest form, the whole process of CT reconstruction can be divided into two parts. Firstly, project the original object or image to obtain the measurement projection

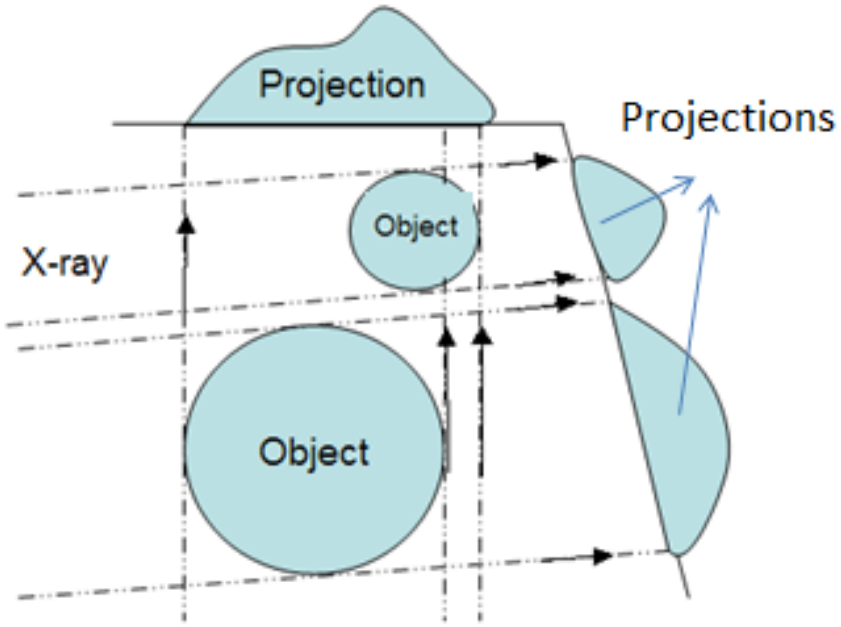
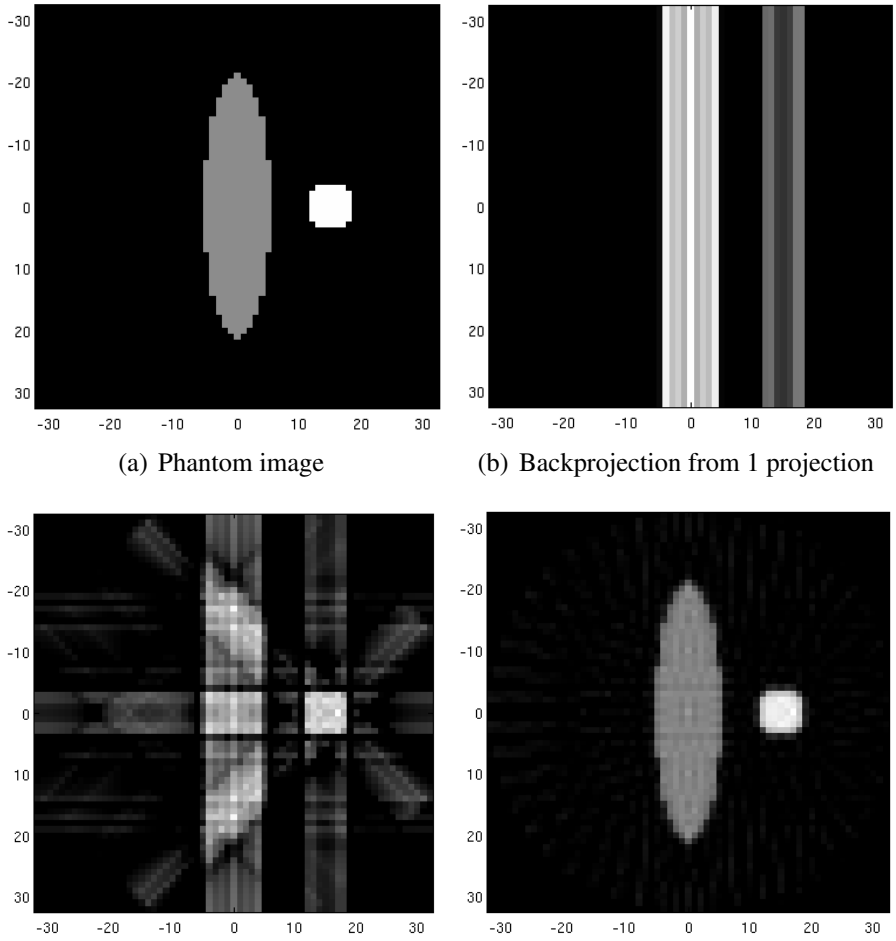


Figure 2.1: Projections from objects.

data. During this process, the object is projected through different angles to get different measurement data. Then to backproject the measurement data to retrieve the object or image. The more measurement data used, the more accurate the reconstruction will be. Figure 2.2 is a computer simulated process of a reconstruction image from its projection data. Figure 2.2(a) is the original image, while the projection data is obtained by the process shown in Figure 2.1 with 32 projections in total while angles are evenly spaced over an angular range from 1 - 180 degrees. Figure 2.2(b) is the reconstruction from projection using 1 set of data (1 angle), while Figure 2.2(c) is reconstructed



(a) Phantom image (b) Backprojection from 1 projection

(c) Backprojection from 4 projections (d) Backprojection from 36 projections

Figure 2.2: Phantom image and its reconstructed images.

from 4 projections (every 45 degree) and Figure 2.2(d) uses 36 projections (every 10 degree).

2.1 Reconstruction methods

Two-dimensional (2D) slice-based CT has been in use for many years, where a number of separated 2D CT slices are acquired to form a '3D' reconstruction. In fact, for this '3D' reconstruction, it is in a simple form that the thin axial slices are stacked one by one on top of each other. In contrast, now more and more research focus on three-dimensional (3D) or volumetric CT, where projection data are reconstructed directly to a 3D form. There are two main kinds of volumetric scanning geometries. One is spiral (helical) and the other is cone-beam type. For spiral geometry, there are a source and a 1D detector array. They are mounted opposite to each other and rotate in a spiral way many times around the patient while for cone-beam geometry, there are a source and a 2D detector array that are mounted opposite to each other and rotate around the patient once as shown in Figure 2.3. The cone-beam geometry is discussed mainly for this project. Examples of visualised projection images of the Rando phantom that we used for experiments in this thesis are shown in Figure 2.4.

In current commercial CBCT reconstruction systems, though many algorithms exist, the filtered back projection (FBP) reconstruction algorithm based on the Feldkamp,

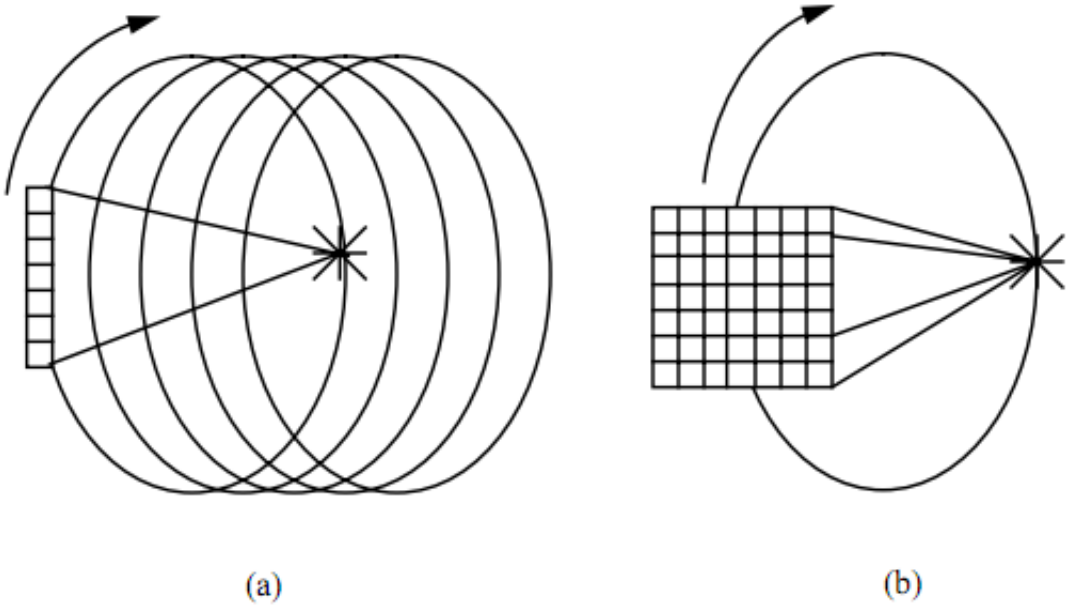


Figure 2.3: Volumetric scanning geometries. (Image was obtained from [1]).

Figure 2.4: Visualised projection images of the Rando phantom

Davis and Kress (FDK) algorithm [8] is still the most frequently used. However, recently, iterative reconstruction algorithms such as expectation-maximization (EM) [10], the algebraic reconstruction technique (ART) [9] and a few others [11, 12] have also been investigated for clinical application [13, 14, 15]. These provide an alternative for commercial tomographic image reconstruction methods.

One problem for CBCT reconstruction is that even where it is possible to access clinical projection data, the lack of a generalised iterative reconstruction package is an obstacle to researchers seeking to optimise or improve reconstructed image characteristics. In addition, many of the previous studies focus on mathematical properties of iterative algorithms, seeking to prove that they will converge (e.g. [16, 17]), but fewer papers describe the actual behaviour using real medical data (e.g. [1, 18] study ART convergence using virtual phantom data).

The multi-instrument data analysis system (MIDAS) is a software package first developed as an image reconstruction software for ionospheric imaging technique [19]. The toolbox is now being extended for medical imaging application in iterative CBCT and the results are verified against commercial FDK results. Projection data are not readily accessible from commercial clinical systems. However, medical physics re-

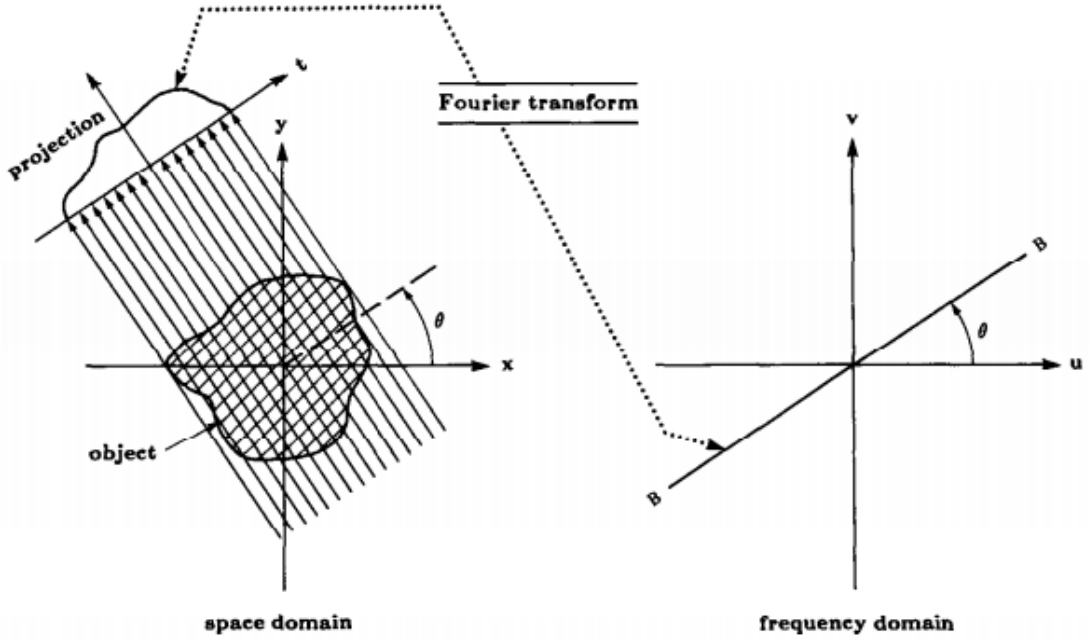


Figure 2.5: Fourier Slice Theorem. Obtained from [2].

searchers at the Christie Hospital have been able to obtain projection sequences, acquired as part of their programme of research into image guided radiotherapy (IGRT) [20, 21].

2.1.1 Analytical approach

There are two main kinds of reconstruction methods. One is the iterative methods which are the main approach used in this thesis. The other one is the analytical reconstruction which is based on Fourier Slice Theorem [4].

Filtered Backprojection (FBP) is one of the most frequently used methods among the analytical type algorithms. Fourier Slice Theorem [4, 22] and Radon transform [3] are applied to the algorithm. The Radon transform is the line integral of a 2D function $f(x, y)$

$$g_\theta(s) = \iint f(x, y)\delta(x\cos\theta + y\sin\theta - s)\, dxdy \tag{2.1}$$

where δ is the impulse response function. The projection image $g_\theta(s)$ is the integral of the object along the line $x\cos\theta + y\sin\theta = s$. Fourier transform (FT) of this projection image $g_\theta(s)$ results a set of sine waves F_θ in the frequency domain, oriented at angle θ and passing the origin as shown in Figure 2.5, which is the main concept of Fourier Slice Theorem. Figure 2.6 is the 2D visualised Rando measurement projection data that are used for experiments in this chapter.

The Fourier Slice Theorem relates the Fourier transform of a projection image from its spatial domain to frequency domain, where the Fourier transform of a projection image at a given angle is a set of sine waves in its frequency domain. Thus, if assume that projections are gathered at enough angles, the image could be reconstructed by simply

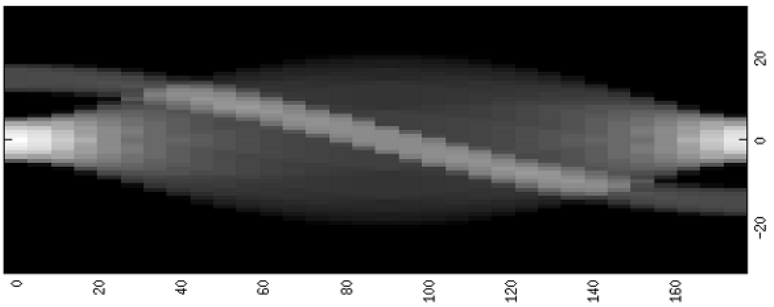


Figure 2.6: Visualised projection image.

using the inverse Fourier transform (IFT). This is an ideal model of reconstruction in tomography, but in practical implementations, a different approach is required, where a filtered version of the projection image is required before the backprojection step.

In almost all applications area of X-ray CT, FBP based methods are used for image reconstruction algorithms. There are several limitations on information from images reconstructed from FBP based methods. They are fast and easy to implement algorithms and most commercial system will provide FBP with some predefine filters.

2.1.2 Iterative approach

In this project, we are implementing iterative reconstruction methods that are capable of handling under sampled data and produces images with higher quality. In the following chapter, we will present implementation and comparative imaging results of

iterative algorithms in volumetric CBCT.

Image reconstruction from computed tomography (CT) commonly uses X-ray data collected evenly spaced over angles ranging from 1 to 360 degrees in full-scan imaging. Half-scan imaging only requires projection from 1 to 180 degrees and in many practical applications, projection data are obtained only within limited angles or limited views [23, 24, 25, 26] and sometimes only limited data (views) are available. Under these conditions, image reconstruction techniques such as FBP may generate artifacts in reconstructed images [27]. A recursive method of iterative reconstruction-reprojection (IRR) was proposed to overcome this problem [24, 28]. The method employs an interpolation operation for the estimation of the missing views, when the projection data are not in a complete form. This was improved later on as projection space iterative reconstruction-reprojection (PSIRR) in [29], where better results were achieved without interpolation, as the backprojection and reprojection process were done in the projection data domain, instead of in the image domain. These algorithms are generally known as analytical methods. There is an alternative reconstruction algorithm which is the class of statistical iterative algorithms [30], where the reconstruction utilises a stochastic model, using maximum likelihood to estimate the parameters. Compared

with analytical reconstruction methods, these algorithms consider the photon statistics in the measurements, resulting in a higher signal-to-noise ratio (SNR) in the reconstructed images [31, 32].

Besides backprojection type algorithms, in CT reconstruction techniques, iterative type algorithms are also considered as an important approach, where many methods were proposed, such as the algebraic reconstruction technique (ART) [9], the simultaneous iterative reconstruction technique (SIRT) [11], the simultaneous algebraic reconstruction technique (SART) [12] and the multiplicative algebraic reconstruction technique (MART) [9]. They were first proposed as an alternative approach for full-scan tomographic image reconstruction. However, improving the reconstructed image quality and reducing the dose of CBCT scans are still the main concerns, especially since CBCT scans are implemented on a daily basis for radiation treatment. In light of these facts, the reduction of radiation dose becomes one of the main clinical concerns. To achieve that, one must either lower the tube energy and/or reduce the number of CT projections. Whenever projection data is insufficient, for example when angles are limited, iterative algorithms have been shown to have a better performance in CBCT reconstruction than algorithms based on FDK [23, 24, 27, 25, 26, 18]. Recently, works

[33, 34, 13] have demonstrated that high resolution prior image information can assist iterative reconstructions where significant amounts of projection data are missing. Among the iterative algorithms, MART is under a different group as the correction is multiplicative rather than additive, and, SART was reported to combine the best of ART and SIRT in [2]. Hence, this thesis studies the behaviour of the algebraic reconstruction technique (ART) [9], the simultaneous algebraic reconstruction technique (SART) [12] and ordered-subset SART (OS-SART) [17] in reference to convergence studies and image quality measurements.

For iterative algorithms, stability, convergence rate and quality of the reconstructed image are important issues. The number of iterations and relaxation parameter are important factors affecting the performance of the algorithm. The conditions for algorithm convergence and appropriate relaxation strategies are discussed in [35]. A detailed study about the convergence of iterative algorithms is presented in [17], but how to evaluate the clinical image is still an open question due to the special requirements of medical applications. Some criteria for reconstructed tomographic image evaluation are considered in [36, 2, 37]. However, there is always a trade-off between stability, computational time and the quality of reconstructed images, as distinct from

consistency with the available projection data, depending on the particular application.

Specifically, in iterative reconstruction algorithms, a relaxation parameter λ is important for image reconstruction because it controls the convergence rate and has the impact of generating/reducing artifacts. Reconstruction algorithms with a larger λ will quickly converge but may produce image degradation, while with a smaller λ, convergence will be slower but often generate smoother images. Obviously, computational cost should also be considered in these scenarios especially in CBCT reconstruction where heavy calculations are required. Therefore, optimisation of the λ-parameter is important and can be achieved by defining different criteria [38, 39]. On the other hand, the quality of CBCT reconstructed images can also be potentially improved by randomising the scanned rays/projections, as has been recently shown in [40, 41], and applied clinically. Referring to [42], the software and algorithms used here are also implemented in the MIDAS platform, but are substantially improved from the previous work reported.

Besides above, the conjugate gradient method, originally proposed by Hestenes and Stiefel [43], is a well-known iterative method for solving sparse systems of equations $Ax = b$ where the matrix A is positive defined [44, 45, 46, 47]. The main problem

with methods like Conjugate gradient least square (CGLS) which involve the normal equations is that the condition number of the matrix $A^\top A$, which usually is the square of the condition number of A. Therefore, for ill-conditioned problems convergence can be slow. However, if the condition number of A is not particularly high, we can reasonably expect CGLS to converge more quickly. Although CGLS is one of the many existing iterative methods, the use for tomography problems is not widespread. We may find examples of its use in limited data problems in [48], in seismic tomography [49, 50], in single-photon emission computed tomography [51] and in a more general CT context [52]. There are some recent developments of CGLS [53, 54], but still not for CBCT reconstruction. In light of these facts, the applications of CGLS in CBCT reconstruction were studied in this thesis. The main challenge in CBCT reconstruction is that matrix A (in $Ax = b$) usually is very large. Storing matrix A requires a large amount of memory. Even supercomputer will fail to do so. This also applies to CGLS algorithm in CBCT reconstruction. By using standard CGLS, only a small amount of the weighting matrix A can be used due to the memory capacity. As a result of insufficient data, the reconstruction usually is very poor. Therefore, a modified CGLS algorithm was proposed to overcome this shortage. It would divide weighting matrix A

into blocks for matrix manipulations. By doing this, A can be loaded blockwise, and, parallel computing can also be applied. Compared with standard CGLS, block CGLS can potentially use a full weighting matrix A for CBCT reconstruction without further enhancing computer standards.

Chapter 3

Iterative reconstructions

The iterative image reconstruction algorithm has potential to improve the image quality compared to traditional FBP based methods. The major advantages of the iterative approach include better noise fidelity and capability to reconstruct the best possible results for down sampled data. It has potential advantages when the projections are not distributed uniformly in angle, or when the projections are sparse. An iterative algorithm starts with an initial image, computes projections from the image, compares the original projection data and updates the image based upon the error between the estimated and the observed projection data.

3.1 MIDAS platform

In this thesis, a study about iterative algorithms in the MIDAS platform has been investigated. In the following, we firstly introduce the data used and the CT scanning system; then we discuss the implementation of the forward projection and time cost; the detail descriptions of algorithms that were used in this project.

3.1.1 The data and system

Several test phantoms were used in this work. In addition, 'Rando' head phantom was also used. The data were provided by North Western Medical Physics at The Christie Hospital in Manchester, which has an image guided radiotherapy research facility equipped with an Elekta Synergy® linac-integrated X-ray cone-beam CT system[1]. This consists of a kilovoltage X-ray tube and flat panel imager mounted onto a radiotherapy linear accelerator gantry at 90 degrees to the megavoltage treatment beam (see Fig.1.1). A 'Rando' anthropomorphic head phantom[2] was scanned to produce 360 X-ray projection images, approximately evenly spaced over an angular range of -100

[1]Elekta Oncology, Crawley, UK.
[2]The Phantom Laboratory, Salem, NY, USA.

to +100 degrees. Images were acquired at 100kV, 10mA and 10ms per projection, with total imaging dose of approximately 1.5mGy. Each projection image contains 512×512 pixels of dimension 1×1mm (note this is down-sampled from the system maximum resolution of 1024×1024). From the system setup, a full scan (360 directions) of a 2D 512×512 image to a 512 size detecter leads to a large matrix Ax (number of scanned direction×number of detecters×image pixels which is $360\times512\times512^2$ double type). Then a full scan (360 directions) of a 3D 512×512×512 image to a 512×512 size detecter will lead to a much larger matrix A ($360 \times 512^2 \times 512^3$ double type).

3.1.2 The forward projection

The geometry of the x-ray source and x-ray detector is defined from source-to-detector distance, source-to-object (centre of rotation) distance and projection angles as supplied with the projection data. Gantry angle dependent corrections to alignment of the projections are also included in these calculations [55]. Transform matrices for conversion from object coordinates (fixed) to projection coordinates (rotating), are then calculated from these values. There is one transform matrix for each projection angle.

A forward model is used with cubic voxel basis functions. When a ray is passes

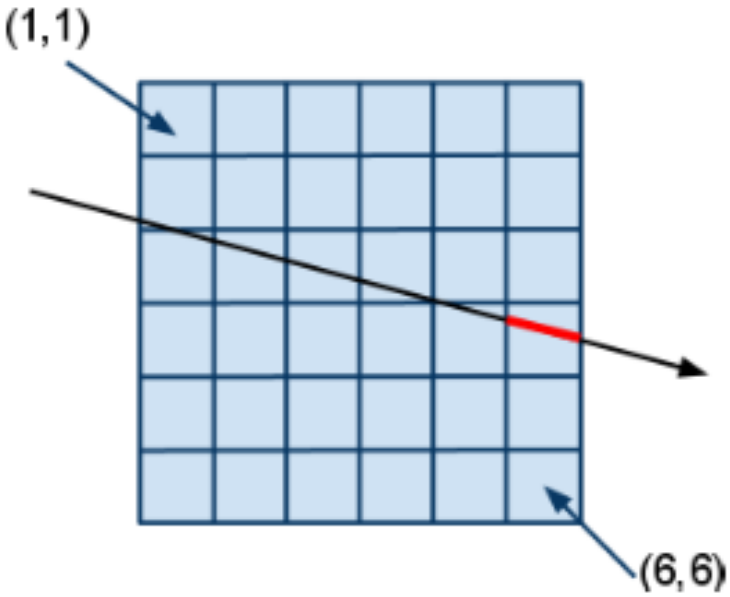


Figure 3.1: Calculating the weighting matrix. The red line is length of the weighting factor for the ray passing through the relevant cell.

through an image cell, the weighting factor is 1 and 0 otherwise. In [56, 57], it has shown that the choice of the basis function used in forward projections strongly influences image quality. Image quality could be dramatically improved by using a more complex forward model.

An image $f(x, y, z)$ is discretised and divided into voxels and each voxel value is a constant number calculated by the average value of the grid, where the total number of voxels is N. When X-ray passes through the voxels from the transmitter to the detector, the measured projection data are obtained by the line integral of the ray sum.

The algebraic equation then can be set up as

$$\sum_{n=1}^{N} A_{mn} x_n = b_m,\ m \in [1, M] \tag{3.1}$$

where A_{mn} is the weighting factor that indicates the contribution of the nth cell in $\boldsymbol{x}$ to the mth ray sum. For convenience, the image value has been assumed constant within each cell. In this project the weight A_{mn} is the length of intersection of the mth ray with the nth cell as shown for a 2D slice in Fig.3.1. In the following experiments, X-rays are assumed linear. The intersections of each ray with the horizontal and vertical planes delineating cells are calculated and hence lengths of intersection are obtained. Our ray tracer operates in both 2D and 3D mode.

The weighting factors play the key point for solving these equations. A simplified version of setting the weighting factors would be: if the ray passes the voxel, set the weighting factor for that voxel to be 1 and 0 vice versa. A more advanced method is to set the factor according to the voxel area covered by the ray or intersection of the ray with the voxel.

3.1.2.1 Projector computational cost

The computational time for projection directions of each iteration are shown in table 3.1. The computer used is 64 bit 3.33GHz Linux with ram of 32GB.

Table 3.1: Time cost for projection directions

1	10	50	100	360
641s	6410s	32050s	64100s	230760s

3.1.3 MIDAS based iterative methods

FDK is performed as a sum of line integral for each projection angle, using nearest-neighbour interpolation, over a defined object grid. The FDK image is then scaled by the number of intersections per volume. As forward projections are done for each projection angle in turn, when corrections are applied in iteratively, they are applied over all rays for a given projection angle at once. The order in which the corrections are applied with respect to projection angle can be sequential or randomised and randomisation is implemented for all presented results here.

Residual norm is defined as $\|x\| = \sqrt{x_1^2 + x_2^2 + ... + x_k^2}, \quad \{ 1 \leq k \leq N \}$, where x is a vector of the difference between the calculated and measured data.

The differences between the iteratively calculated and measured projection data are used for comparison in this work. However, a good residual norm result may not always indicate a good reconstructed image, therefore comparison should also be made between images. In this work, besides norm differences comparison, image 1D and 2D profiles are used, which gives indications of the contrast between the bones and soft tissues and presents grey scale data extracted from the reconstructed images respectively.

In general, due to noise and discrete sampling of data, the relaxation parameter is vital to the computation. If λ is not properly chosen, $f^{(i)}$ will diverge from f after just a few iterations. Even with a well-chosen value of λ, there is still be the possibility that, with increasing iterations, $f^{(i)}$ will initially converge and then subsequently diverge. The problem then is to set both an appropriate λ and number of iterations.

Norm differences could be used in this case to evaluate the setting of the relaxation parameter λ, taking the number of iterations and the quality of the reconstructed image into account. The whole process of the reconstruction can be divided into two stages, reprojection and reconstruction. For example, as ART progresses, the image $f^{(i)}$ in the ith iteration will be different from the image $f^{(i+1)}$ in the $(i + 1)$th iteration. During

each iteration, ART generates a new set of calculated projection data $p^{(i)}$ from $f^{(i)}$. As the reconstructed image evolves, the differences between calculated projection data $p^{(i)}$ and original projection data p should decrease. If

$$E = \|p^{(i)} - p\| \tag{3.2}$$

then

$$\Delta E = E^{(i+1)} - E^{i} \tag{3.3}$$

where ΔE is the projection mismatch errors and it shows the convergence or divergence of a given ART reconstruction. It is assumed that the number of iterations needed to obtain better image quality can be determined by the value of i when E is minimum. Hence, Δf is the differences between the iterative FDK images.

The MIDAS based iterative algorithm used for this experiment is

$$f^{(i+1)} = f^{(i)} + \lambda \Delta f^{(i)} \tag{3.4}$$

where λ is the relaxation parameter, which controls the convergence rate. $f(x, y, z)$ are

the image values; i is the number of iterations. Including a relaxation parameter can improve the quality of the reconstructions, but usually at the expense of the rate of convergence. Depending on the application, different strategies are applied for choosing the most appropriate relaxation parameter and other settings.

3.1.3.1 Computational time

The computational time for each MIDAS based iterative iteration is shown in table 3.2. The computer used is 64 bit 3.33GHz Linux with ram of 32GB.

For each iteration with full data set, around 60% of the time is on forward projection and 30% of the time on backward projection.

Table 3.2: Time cost for one ART iteration with available data

Full data	1/2 data	1/3 data	1/4 data	1/5 data
1600s	800s	540s	400s	320s

3.2 Iterative methods

For iterative methods, no filtering is applied to the data. A forward projection is carried out for each transform matrix in turn, summing along raypaths as defined by objects

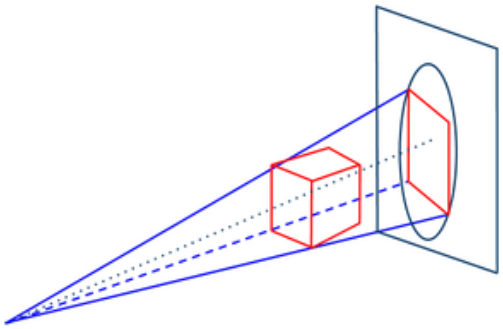


Figure 3.2: An object is scanned by CBCT and projection data is obtained from a flat panel detector.

and detector geometry. As with the forward projection, nearest-neighbour interpolation is used.

The data used for the image reconstruction are the 2D projections of the CBCT scan. An example of the acquisition of a 2D projection is sketched in Fig.3.2. There are mainly two types of approaches that are used for CBCT image reconstruction. One of them is the filtered backprojection (FBP) type reconstruction, which is based on the Fourier Slice Theorem [4]. Besides that, an entirely different type of algorithm is the algebraic type algorithms. Algebraic methods define a linear system of equations and then propose a method to solve them. The tomographic image consists of an array of unknowns, and then linear algebraic equations are set up to calculate the unknown values in the array according to weighting factors and projection data. Iterative meth-

ods have been widely studied and investigated for clinical application [13, 14, 15]. They offer the potential to generate high quality images and are known to be advantageous for undersampled projection data. There are several algebraic reconstruction algorithms that are well investigated for image reconstruction, such as the algebraic reconstruction technique (ART) [9], the simultaneous algebraic reconstruction technique (SART) [12], and ordered-subset SART (OS-SART).

3.2.1 Traditional iterative methods

In the section, traditional iterative methods that were used in this project are presented in details.

3.2.1.1 ART, SART and OS-SART

In this section, ART, SART and OS-SART are summarised based on the definitions in [17]. These methods solve the classic linear algebraic system of equations $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$. If $\exists \boldsymbol{x} : \mathbf{A}\boldsymbol{x} = \boldsymbol{b}$, then the problem is consistent, else inconsistent. Our artificial problem in Sec.6.1.2 is an example of a consistent problem; the reconstruction of the undersampled 'RANDO' phantom is an example of an inconsistent problem. Most real

world problems are inconsistent due to discretisation and measurement errors. First let

$$\boldsymbol{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_M \end{bmatrix} \in \mathbb{R}^{M \times 1} \tag{3.5}$$

In the block iterative case, $\boldsymbol{b}$ can be divided into B blocks so

$$\boldsymbol{b} = \begin{bmatrix} \boldsymbol{b}(1) \\ \vdots \\ \boldsymbol{b}(B) \end{bmatrix} \tag{3.6}$$

where $\boldsymbol{b}(\beta) \in \mathbb{R}^{M(\beta) \times 1}$ for block index $\beta \in [1, B]$. Next let

$$\mathbf{A} = \begin{bmatrix} \tilde{\boldsymbol{a}}_1^\top \\ \vdots \\ \tilde{\boldsymbol{a}}_M^\top \end{bmatrix} = [\boldsymbol{a}_1 \cdots \boldsymbol{a}_N] \in \mathbb{R}^{M \times N} \tag{3.7}$$

where $\tilde{\boldsymbol{a}}_M^\top$ is the Mth row of $\mathbf{A}$, and $\boldsymbol{a}_N$ is the Nth column of $\mathbf{A}$.

In the block iterative case $\mathbf{A}$ can be subdivided into B partitions,

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}(1) \\ \vdots \\ \mathbf{A}(B) \end{bmatrix} \tag{3.8}$$

such that, for $\beta \in [1, B]$

$$\mathbf{A}(\beta) = \begin{bmatrix} \tilde{\boldsymbol{a}}_1(\beta)^\top \\ \vdots \\ \tilde{\boldsymbol{a}}_{M(\beta)}(\beta)^\top \end{bmatrix} = [\boldsymbol{a}_1(\beta) \cdots \boldsymbol{a}_N(\beta)] \in \mathbb{R}^{M(\beta) \times N} \tag{3.9}$$

Then the system of equations can be solved by the following iterative techniques.

Note that the equations are constrained by iteration $k \geq 0$. For ART,

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \lambda_n \frac{1}{\|\tilde{\boldsymbol{a}}_i\|_{\mathrm{L2}}^2} \tilde{\boldsymbol{a}}_i (b_i - \tilde{\boldsymbol{a}}_i^\top \boldsymbol{x}^{(k)}) \tag{3.10}$$

where i is the row index and $i \in (k \, mod \, M) + 1$ and $\|\tilde{\boldsymbol{a}}_i\|_{\mathrm{L2}}^2 = \sum_{j=1}^{N} (\mathbf{A}_{ij})^2$. Then SART can be written as,

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \lambda_n \mathbf{V}^{-1} \mathbf{A}^\top \mathbf{W} (\boldsymbol{b} - \mathbf{A}\boldsymbol{x}^{(k)}) \tag{3.11}$$

where $\mathbf{V} \in \mathbb{R}^{N \times N}$ and $\mathbf{W} \in \mathbb{R}^{M \times M}$ are both diagonal and

$$\mathbf{V} = \text{diag}\left\{ \|\boldsymbol{a}_j\|_{\text{L1}} : j \in [1, N] \right\} \tag{3.12}$$

$$\mathbf{W}^{-1} = \text{diag}\left\{ \|\tilde{\boldsymbol{a}}_i\|_{\text{L1}} : i \in [1, M] \right\} \tag{3.13}$$

where

$$\|\boldsymbol{a}_j\|_{\text{L1}} = \sum_{i=1}^{M} |A_{ij}| \tag{3.14}$$

$$\|\tilde{\boldsymbol{a}}_i\|_{\text{L1}} = \sum_{j=1}^{N} |A_{ij}| \tag{3.15}$$

The diag operator simply maps the elements of a column vector into the leading diagonal of a square matrix, where off-diagonal elements are set to zero. Next, OS-SART can be written as,

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \lambda_n \mathbf{V}^{-1} [\mathbf{A}(\beta)]^{\top} [\mathbf{W}(\beta)] ([\boldsymbol{b}(\beta)] - [\mathbf{A}(\beta)] \boldsymbol{x}^{(k)}) \tag{3.16}$$

where $\mathbf{W}(\beta) \in \mathbb{R}^{M(\beta) \times M(\beta)}$ is diagonal and

$$[\mathbf{W}(\beta)]^{-1} = \text{diag}\,\{||\tilde{\boldsymbol{a}}_i(\beta)||_{\text{L1}} : i \in [1, M(\beta)]\} \tag{3.17}$$

In effect these algorithms solve weighted least squares problems. Hence $\lim_{k\to\infty} \boldsymbol{x}^k = \hat{\boldsymbol{x}}$ where $\hat{\boldsymbol{x}} = \arg_{\boldsymbol{x}} \min f$ and $f = (\mathbf{A}\mathbf{x} - \boldsymbol{b})^\top \mathbf{W}(\mathbf{A}\mathbf{x} - \boldsymbol{b})$. For SART and OS-SART, $\mathbf{W}$ is given as in Eq.3.13; for ART it is given as $\mathbf{W}^{-1} = \text{diag}\,\{||\tilde{\boldsymbol{a}}_i||^2_{\text{L2}}\}$. Since $\mathbf{W}$ differs for ART, SART/OS-SART, these 3 algorithms solve two different weighted least squares problems and the images obtained will differ if the problem is inconsistent.

As explained in Sec.6.1.1, the measured data in $\boldsymbol{b}$ is collected from 90 projection angles. With 512 measurements per angle, the data is arranged in scan order. Hence if $(M(\beta)) = 512$, one block corresponds to one projection. For OS-SART, convergence of data is suboptimal since the blocks are not well balanced.

3.2.1.2 Choosing the relaxation schedule

The relaxation parameter λ controls the convergence rate. Similar to [58], λ is varied according to a schedule $\lambda_n, n \geq 1$ where n is the cycle index and

$$\lambda_n = \sigma \frac{1}{n^\alpha} \;:\; \lambda_n \in (0, 2) \tag{3.18}$$

and σ is the scaling factor. Note that λ is constructed in this way so that the iterative algorithms in Sec.3.2.1 are guaranteed to converge [17]. Additionally $\boldsymbol{x}^{(0)} = \mathbf{0}$ so that the iterative algorithms converge in some weighted least squares sense. Here a cycle is defined as one complete pass through all the projection data, where each ray path is used strictly once. When $\alpha = 0$, $\lambda_n = \sigma \, \forall n$, which is one of the common schedules chosen for iterative algorithms, but does not guarantee convergence for an inconsistent problem (ie $\forall \boldsymbol{x}, \mathrm{A}\boldsymbol{x} \neq \boldsymbol{b}$). In this work, comparisons of different relaxation schedules were investigated and results will be presented in the following chapters.

3.2.2 CG and LSQR

The CBCT reconstruction algorithm essentially discretises the scanning process (forward projection) into a system of linear equations, which can then be solved in a (possibly regularized) linear least squares sense to reconstruct a scanned image from projection data. The resulting system of equations is typically much too large to solve using a direct method. However, as the matrix A is usually very sparse, iterative methods can be used to compute a solution.

The conjugate gradients (CG) method is a widely used iterative algorithm for the

numerical solution of large sparse systems of linear equations, whose matrix is symmetric and positive-definite. An overview of applications of CG, and generalizations to indefinite or non-symmetric matrices, can be found, for example, in [47].

LSQR is a variant of the conjugate gradient method which can be used for solving non-symmetric linear equations $Ax = b$ and least squares problems $\min_x \|b - Ax\|_2$. If A has full column rank (linearly independent columns), $A^\top A$ is positive-definite and the conjugate gradient algorithm can be applied to the normal equations $A^\top Ax = A^\top b$. The resulting algorithm can be implemented in several ways. In this thesis, the algorithm LSQR has been used, which is mathematically equivalent to CG on the normal equations but has favorable properties in floating-point arithmetic [59].

3.2.3 Blockwise matrix vector multiplication

CG-type methods such as LSQR do not require access to the full matrix A. All that is required is one matrix-vector multiplication with each of A and $A^\top$ per iteration. Therefore, blockwise matrix-vector multiplication within LSQR provides a solution to memory capacity constraints. Even given a extremely large 3D weighting factor matrix A (e.g. 360×512^2 by 256^3 in our case), the algorithm has the ability to handle

it, as it does not require the matrix A to be loaded into memory all at once. In addition, the matrix-vector multiplications can be implemented to run in parallel, which may increase the reconstruction speed dramatically.

First, the data are divided into blocks according to current computer memory capacity. Recall that the matrix A resulting from a 3D full scan has dimensions $M = 360 \times 512^2$ by $N = 256^3$. The matrix A can be divided into, for example, 20 block rows,

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_{20} \end{bmatrix}, \tag{3.19}$$

where the index j in A_j refers to the position of each block within the matrix A. Each block A_j contains data from 18 projection directions, and is itself a matrix of dimension $M/20$ by N. Matrix-vector multiplications can then easily be implemented blockwise.

If v is a vector of length N,

$$Av = \begin{bmatrix} A_1 v \\ A_2 v \\ \vdots \\ A_{20} v \end{bmatrix},$$

where each $A_j v$ involves a matrix-vector multiplication with the block A_j. Similarly, if u is a vector of length M, it can be partitioned into blocks

$$u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{20} \end{bmatrix},$$

where each u_j is a vector of length $M/20$, and

$$A^\top u = \sum_{j=1}^{20} A_j^\top u_j.$$

Thus, each block weighting matrix A_j can be loaded on the fly during the matrix-vector

multiplication. This makes it possible to run LSQR using the full weighting matrix A.

3.2.4 Regularization

If the matrix A is very ill-conditioned or if the the data A and b contain large measurement or discretization errors, the LS solution may be dominated by noise and may not be physically meaningful. Several techniques can be used to regularize the computed solution. (See for example the overview in [46].) For instance, terminating LSQR after very few iterations is a form of regularization. In Tikhonov regularization, instead of minimizing the residual norm $\|b - Ax\|_2$, one solves $\min_x \|b - Ax\|_2^2 + \alpha^2\|x\|_2^2$, equivalently,

$$\min_x \left\| \begin{bmatrix} b \\ 0 \end{bmatrix} - \begin{bmatrix} A \\ \alpha I \end{bmatrix} x \right\|_2, \tag{3.20}$$

where I is the identity matrix of size N and α is a suitably chosen regularization parameter.

Tikhonov regularization can be implemented in LSQR at virtually no extra cost (see [60]). An equivalent (but slightly less efficient) approach is to treat the scaled identity matrix in (3.20) as one extra block during the matrix vector multiplications.

For example, in (3.19), the matrix αI would be block 21, or A_{21}.

Results using both un-regularized and regularized LSQR are presented in the next section.

3.3 Conclusion

In this chapter, iterative algorithms for CBCT reconstruction that were used in this project are introduced. Results and discussions will then be presented in the following chapters.

Chapter 4

Image quality assessment

The study concentrates on the optimisation of the relaxation parameter and number of iterations for the criteria of stability, convergence rate, quality of reconstructed image and recovery of sharp edges for full and limited data iterative reconstruction. For the latter, testing is performed using 1/2, 1/3, 1/4 and 1/5 of X-ray projections angularly sampled in an even manner from the full projection set.

Using the full 360 projection data set a 3D reconstruction of 256×256×256 voxels with resolution 1mm in each direction was produced using iterative techniques. Other volume reconstructions were performed using 180, 120, 90 and 72 projections. A

'reference image' was reconstructed using the COBRA cone beam software developers package from EXXIM[1]. This contains an implementation of FDK FBP, which is a useful benchmark for the iterative techniques described in this work.

The criteria about how to evaluate the clinical image is still an open question due to the special requirements of medical applications. Some criteria are mentioned in [36, 2, 37]. Different criteria to evaluate the clinical image are generated for different purposes, as there is always a trade-off among them. Besides the optimisation measures mentioned in previous chapter, there are also other image quality measurements. Depending on the imaging requirements and the applications, different approaches are selected for image quality evaluation. Here we study some of these measures that will be used in this project.

4.1 Uniformity

In Figure 4.1, four regions of interest (ROI) are defined within skull on slice 128 (central slice). The phantom is made of a uniform density material, which is approximately tissue equivalent. Hence, each of these regions would have the same mean pixel value

[1]Exxim Computing Corporation, Pleasanton, CA, USA.

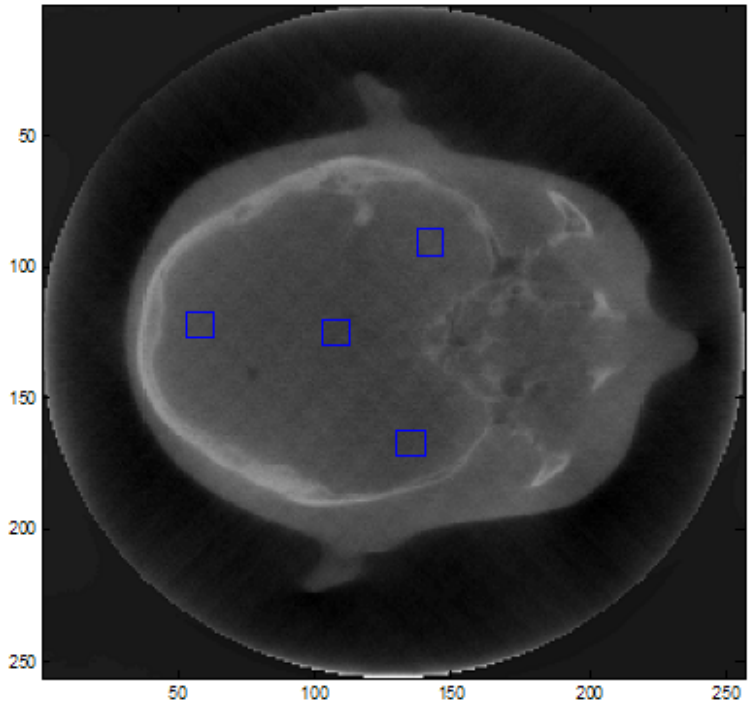


Figure 4.1: Uniformity sample.

for a good quality image. The mean value for each ROI was calculated and uniformity was defined as the maximum difference between any two ROIs quoted as a percentage of the central region mean value. Note that the uniformity value will be affected by scatter artefacts (causing cupping - increased pixel values towards the edge of the object), which are not taken into account in the reconstruction process.

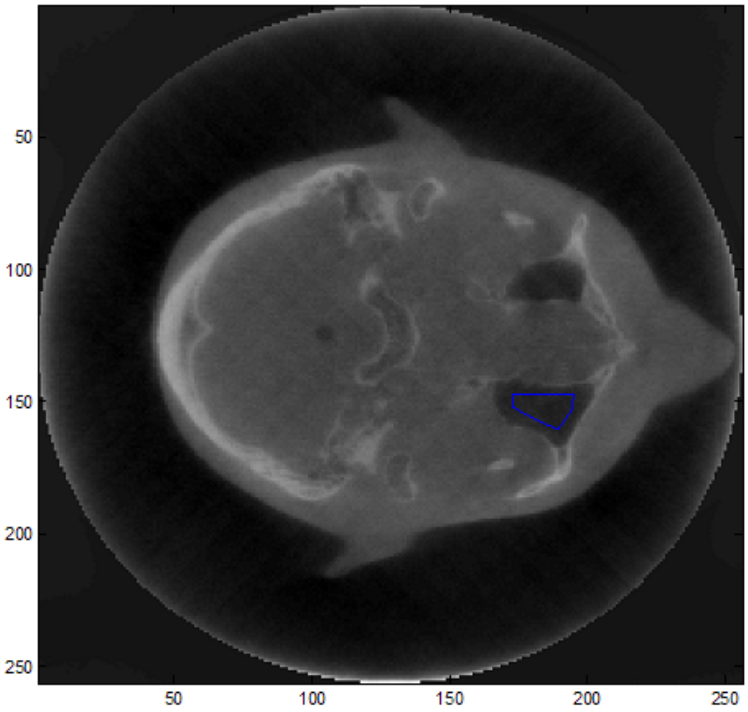


Figure 4.2: Contrast sample.

4.2 Noise

The average of the standard deviations of the four regions of interest on slice 128 was used as a measure of image noise.

4.3 Contrast

In Figure 4.2, an air cavity within the phantom on slice 150 was used to measure image contrast. Contrast was defined as the difference between the mean soft tissue

pixel value (taken from the mean value of the central ROI on slice 128) and the mean air pixel value (taken as the mean pixel value within the cavity on slice 150). Note that the absolute value of contrast does not necessarily correlate with how good the image is, since the overall normalisation of the images may vary.

4.4 Contrast to Noise Ratio

Contrast to Noise Ratio (CNR) is used as a measure of object detectability which is less sensitive to overall image normalisation. Here CNR was defined as contrast/noise, with the contrast and noise parameters as defined above.

4.5 Error measurements

The error measurements that are used in Sec.6.1 are defined as

$$\text{RMS projection error} = (\boldsymbol{b} - \mathbf{A}\mathbf{x})^{\top}(\boldsymbol{b} - \mathbf{A}\mathbf{x}) \tag{4.1}$$

$$\text{L1 reconstruction error} = \|\boldsymbol{x}\|_{\text{L1}} \tag{4.2}$$

$$\text{L2 reconstruction error} = \|\boldsymbol{x}\|_{\text{L2}} \tag{4.3}$$

For simulated data in Sec.6.1.2, the RMS reconstruction error is defined as,

$$f(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{x}')^{\top}\mathbf{V}(\boldsymbol{x} - \boldsymbol{x}') \tag{4.4}$$

where $\boldsymbol{x}'$ is the truth. The RMS reconstruction error is calculated without a mask. For the 'RANDO' data in Sec.6.1.3 and 6.1.5,

$$f(\boldsymbol{x}) = (\frac{\boldsymbol{x}}{s} - \frac{\boldsymbol{x}'}{s'})^{\top}\mathbf{V}(\frac{\boldsymbol{x}}{s} - \frac{\boldsymbol{x}'}{s'}) \tag{4.5}$$

where $\boldsymbol{x}'$ is the FDK reconstruction. Since the FDK and iterative reconstruction are in a different dynamic range, scaling is required. Here $s' = \frac{1}{N}\sum_{i=1}^{N} x_i'$ and $s = 44.9313$, which is the mean value of $\boldsymbol{x}$ taken from the last reconstruction of OS-SART with one full projection for each block ($\sigma = 1, \alpha = 0.1$). It is considered to be a good reconstruction based on the RMS projection error defined. Note that for ART, $\mathbf{V} = \mathbf{I}$.

$$P = \sum_{a=1}^{N} f(a, b, c), \quad \{\, b \in Z, c \in Z \mid 1 \le b \le N, 1 \le c \le N \,\} \tag{4.6}$$

$$P = \sum_{a=1}^{N} \sum_{b=1}^{N} f(a, b, c), \quad \{ c \in Z \mid 1 \leq c \leq N \} \tag{4.7}$$

Equation 4.6 indicates the 1D profiles of the reconstructed image, highlighting the attenuation of X-ray crossing different tissues, while Equation 4.7 is the transverse plane 2D profiles. The quality of the reconstructed image is assessed by comparison to a reference image, which is assumed to be of high quality. The difference between the reconstructed and reference image is defined as

$$\Delta P = \sigma P_i / P_{ref} \tag{4.8}$$

where σ is the scaling factor and i is the number of iteration. When the reconstruction converges, ΔP becomes smaller as the optimal image quality is achieved. In addition to this, root mean square errors between the reconstructed and reference images are used as well.

The ideal number of iterations and optimal relaxation parameter will be found when both ΔE and ΔP are minimised.

4.6 Image quality measurements

For the purpose of evaluation, besides the convergence property, image quality parameters (uniformity and noise) are also introduced for experimental data. As shown in Fig.4.1, four regions of interest can be defined within the skull on axial slice 128 (central slice). The phantom is made of a uniform density material, which is approximately tissue equivalent. Hence, each of these regions are expected to have the same mean pixel value for a good quality image as mentioned before. The mean value for each ROI was calculated and uniformity was defined as the maximum difference between any two ROIs quoted as a percentage of the central region mean value. In addition, the average of the standard deviations of the four regions of interest on axial slice 128 was used as a measure of image noise. By combining the convergence study and image quality parameters, a simple function can be derived as,

$$h(\mathbf{x};\mu) = (1-\mu)\frac{f(\mathbf{x})}{\bar{f}(\mathbf{x})} + \mu\frac{g(\mathbf{x})}{\bar{g}(\mathbf{x})} \tag{4.9}$$

where $f(\mathbf{x})$ is an indication of convergence of the reconstruction, $g(\mathbf{x})$ is an image quality parameter and $\mu \in [0, 1]$, $\bar{f}(\mathbf{x})$ is the mean of $f(\mathbf{x})$ and $\bar{g}(\mathbf{x})$ is the mean of $g(\mathbf{x})$. For

example, $f(\mathbf{x})$ can be taken as the root mean square reconstruction error and $g(\mathbf{x})$ is one of the above uniformity or noise parameters. If $\mu = 0$, $h(\mathbf{x}) = f(\mathbf{x})/\bar{f}(\mathbf{x})$, which means that reconstruction is considered according to the quality of convergence solely. However for $\mu = 1$, $h(\mathbf{x}) = g(\mathbf{x})/\bar{g}(\mathbf{x})$ indicating that reconstruction is considered according to reconstructed image quality solely. Note that if

$$\frac{\partial h(\mathbf{x};\mu)}{\partial \mu} = 0 \tag{4.10}$$

then

$$-\frac{f(\mathbf{x})}{\bar{f}(\mathbf{x})} + \frac{g(\mathbf{x})}{\bar{g}(\mathbf{x})} = 0 \tag{4.11}$$

Hence if $\exists \tilde{\boldsymbol{x}}$:

$$\frac{f(\tilde{\boldsymbol{x}})}{\bar{f}(\tilde{\boldsymbol{x}})} = \frac{g(\tilde{\boldsymbol{x}})}{\bar{g}(\tilde{\boldsymbol{x}})} \tag{4.12}$$

then there is a stationary point in h as μ varies at that $\tilde{\boldsymbol{x}}$. Ideally the iterative algorithms are run until $h(\boldsymbol{x};\mu)$ is minimised. Note however that the function may have local minima.

4.7 Conclusion

In this chapter, we presented some quantitative image quality measures that are used to evaluate the performance of the image reconstruction algorithms. There are several more image quality measures that could be used. We will use these implemented techniques throughout this project for assessing image quality. Please note that the image quality measures introduced in this chapter are a standard package that medical physicists will use in evaluation. Some of the measures may not be used in this thesis but for a background presentation purpose only.

Chapter 5

MIDAS based iterative methods

Cone beam computed tomography (CBCT) is widely used in many real life applications and it enables a volumetric image reconstruction from a set of 2D projection data. It plays an important role in image guided radiation therapy (IGRT). Although CBCT reconstruction are studied widely, the lack of a generalised iterative reconstruction package prevents researchers from investigating the reconstructed algorithms for further improvements. MIDAS is a software package first developed as an image reconstruction software for ionospheric imaging technique. The toolbox is now being extended for medical imaging application in iterative CBCT. In this chapter, iterative

methods are studied on the MIDAS platform. This provides a generalised ready-to-use platform with iterative methods available, allowing scientific researchers to investigate the behaviour of iterative methods in a more general context.

5.1 Results and discussion

To be specific, this chapter studies the convergence (or divergence) with different relaxation parameters and image quality at each iteration. The convergence of the MIDAS based iterative reconstruction is considered in terms of residual norm and visual image quality. This is supplemented by profiles of grey scale data extracted along the rows, transverse plane of reconstructed images and root mean square errors between images, facilitating visual comparison with the FDK reference image. Initially we test the algorithms using simulated and measured data. Figure 5.1 shows the FDK reference image generated from measured projections.

For this study, we have chosen relaxation parameters λ within a certain range. As λ values higher than this range will course strong artifacts in the first few iterations while λ values lower than this range will have less artifacts but has a slow convergence.

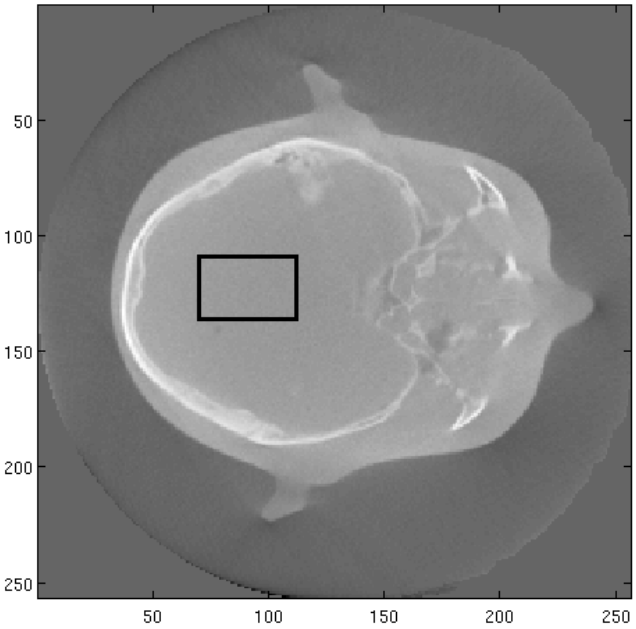
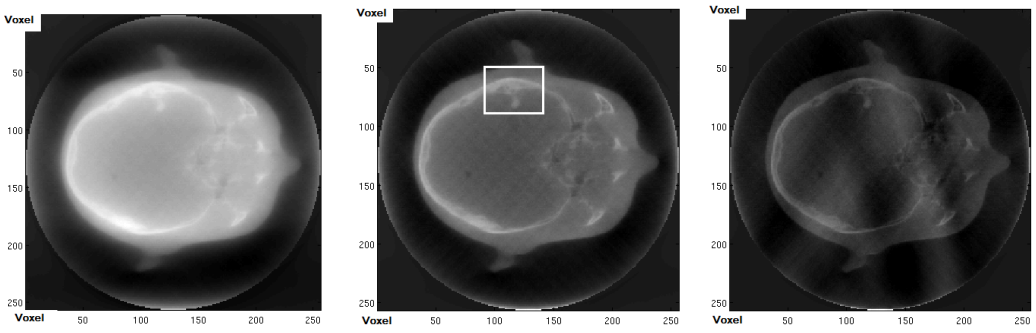
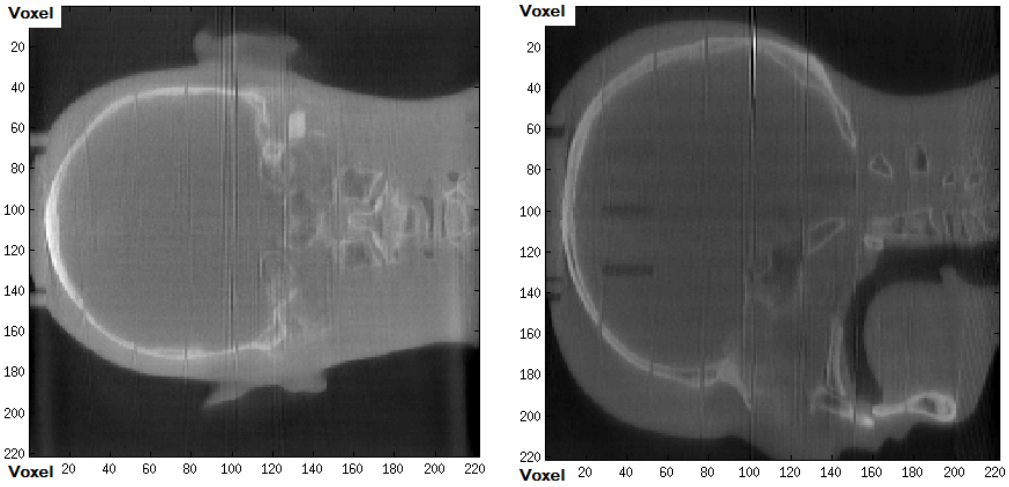


Figure 5.1: Phantom reconstructed image using full data set and FDK algorithm is used as a reference image.



(a) 1th iteration (b) 5th iteration (c) 15th iteration

Figure 5.2: Reconstruction with full data at λ of 0.0146.



(a) Coronal (b) Sagittal

Figure 5.3: Reconstruction with full data at λ of 0.0146 for different directions.

In order to test the convergence rates of the MIDAS based iterative reconstruction, an incremental reduction of the relaxation parameter was implemented, varying from 0.02 to 0.002 in steps of 0.0018. With the full projection data set a maximum of 15 iterations was used, since initial results showed that further iterations caused divergence of the algorithm. However, with lower values of λ the reconstruction takes longer to diverge, so with down sampled projection data sets (1/2, 1/3, 1/4 and 1/5), a maximum of 20 iterations was used.

Figure 5.2 presents the reconstructed images at $\lambda = 0.0146$ using the full data set, and Figure 5.3 shows coronal and sagittal views of Figure 5.2(b). Figure 5.4 shows the result of mismatch errors after each of the 15 iterations for λ of 0.02, 0.0146, 0.0074

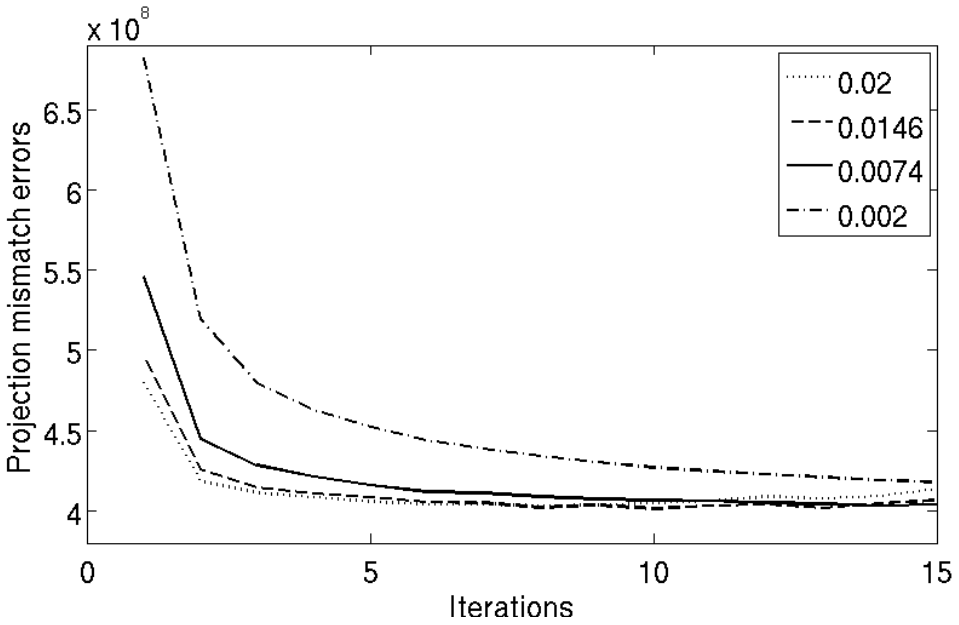
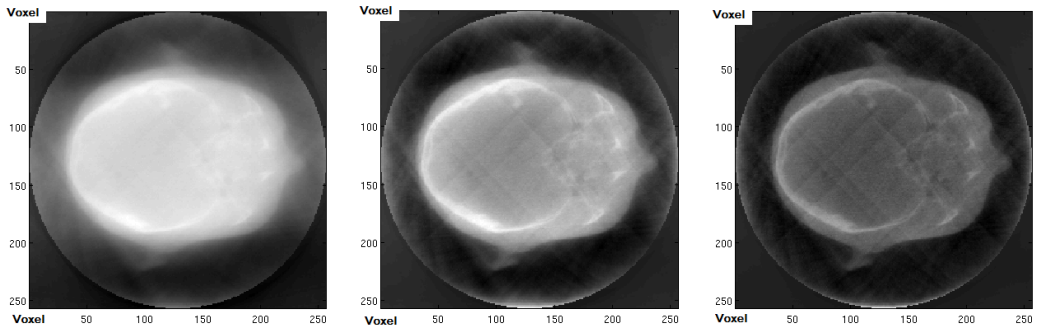


Figure 5.4: Convergence for λ of 0.02, 0.0146, 0.0074 and 0.002 for full data.

and 0.002 respectively. Note that comparisons were started after the first iteration, which causes the initial mismatching errors to vary. From the figures, it can be seen that the higher the relaxation parameter, the quicker the convergence. Compared with 0.002, a relaxation parameter of 0.02 reaches a relatively low level of norm differences quickly, after only 5 iterations. However, even after 15 iterations, a relaxation parameter 0.002 still does not achieve the same level of norm differences. For higher values of λ when the number of iterations increases, there may be fluctuations in image quality as well as divergence of the algorithm. A λ of 0.002 performs better in this case; a



(a) 1th iteration (b) 5th iteration (c) 15th iteration

Figure 5.5: Reconstruction with one third of data at λ of 0.0146.

smoother approach to convergence is expected and the norm differences improve with increased iterations. From Figure 5.4, the stability and convergence rate for λ can be obtained.

When down sampled measured projection data (like 1/2 and 1/3) are used in iterative reconstruction, the reconstructions also smoothly converge as the number of iterations increases. For such down sampled projection reconstructions, differences in the mismatch errors (Equation 3.2) can be readily observed when comparing Figures 5.4 and 5.6. Figure 5.6 indicate the convergence results for different relaxation parameters reconstructed from 1/3 of the full projection data. When reduced data sets are used, the reconstructions converge for all values of λ over the range of iterations used. The plots for λ of 0.02, 0.0146 and 0.002 all show similar trends to those in Figure 5.4.

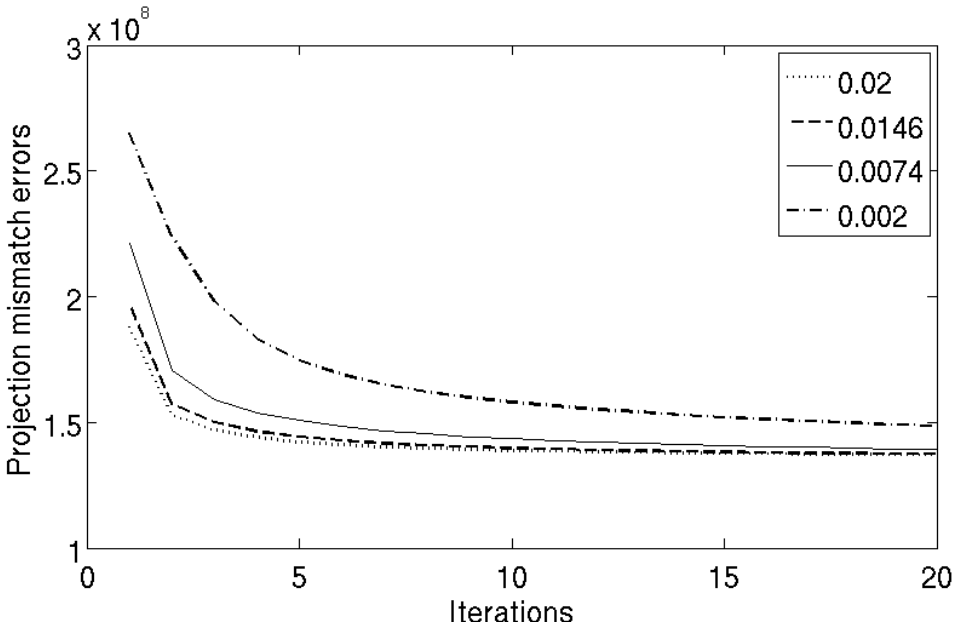


Figure 5.6: Convergence for λ of 0.02, 0.0146, 0.0074 and 0.002 for one third of data. For a higher λ, convergence has improved compared to that seen in Figure 5.4 as with less data are contribute to the reconstruction, the process can endure longer iterations until it diverges.

From Figure 5.4 and 5.6, norm differences are demonstrated as an effective way of indicating convergence, not only for a full projection set but also for down-sampling data. However, norm differences calculated from iterated projections cannot be considered as the only criteria for convergence. Figure 5.4 shows the algorithm converges for $\lambda = 0.0146$ as iteration increases. However, the quality of reconstructed image may

not in its good condition. (e.g. Figure 5.2(c) shows a reconstruction where the algorithm diverged). In [61, 62, 63], it was shown that in helical CT scanning, the object being scanned could enter and leave the X-ray cone, which can lead to artifacts in the reconstructed images. This problem can be reduced by adding an aperture weighting function for each projection of p_k in the back projection formula. Possibly due to the presence of artifacts, according to our study, empirically, reconstruction will eventually fluctuate or diverge. However, depending on the selection of λ and availability of projection data, the rate of divergence of reconstruction varies. The lower the relaxation parameter and/or the lower the availability of projection data, the larger number of iterations required for the process to diverge. Of course, with lower relaxation parameter, the convergence is time consuming while with less availability of projection data, other artifacts such as streaks shown in Figure 5.5(c) may appear.

Besides projection differences, the norm of differences between the reference and reconstructed image is generated for this phantom study. Figure 5.8 shows 1D plots across the central row of the images calculated using one third of the available data, compared to the same row from a commercial FDK image. The first iteration shows little contrast across the image, but by the tenth iteration the edges have been recovered

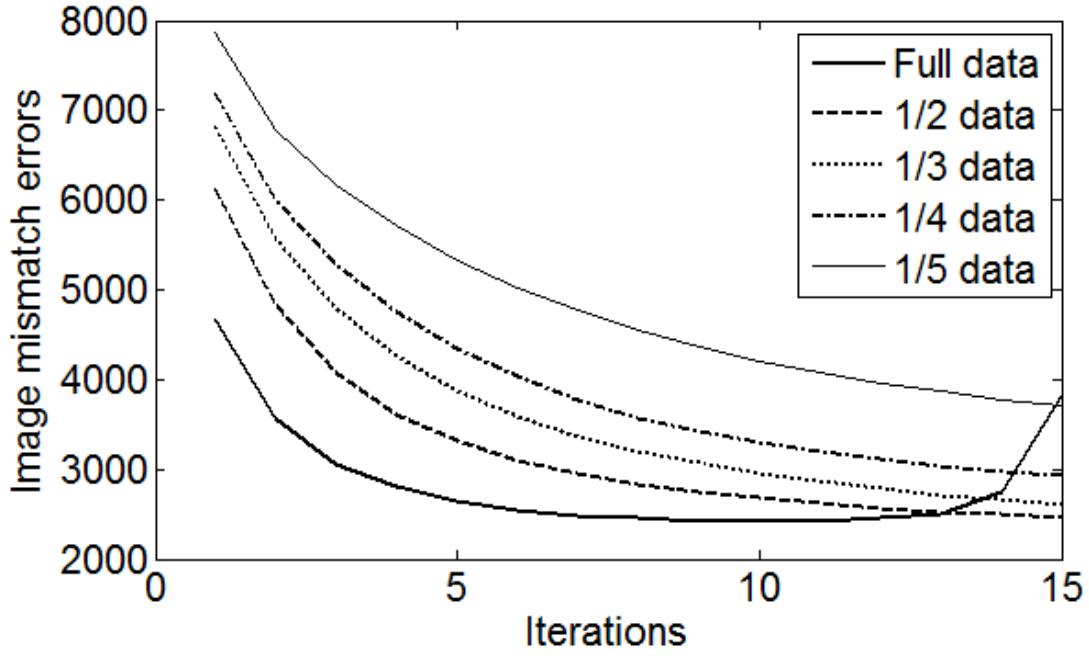


Figure 5.7: Image transverse plots for different data set at $f(x = 120 : 150, y = 55 : 78, z = 128)$.

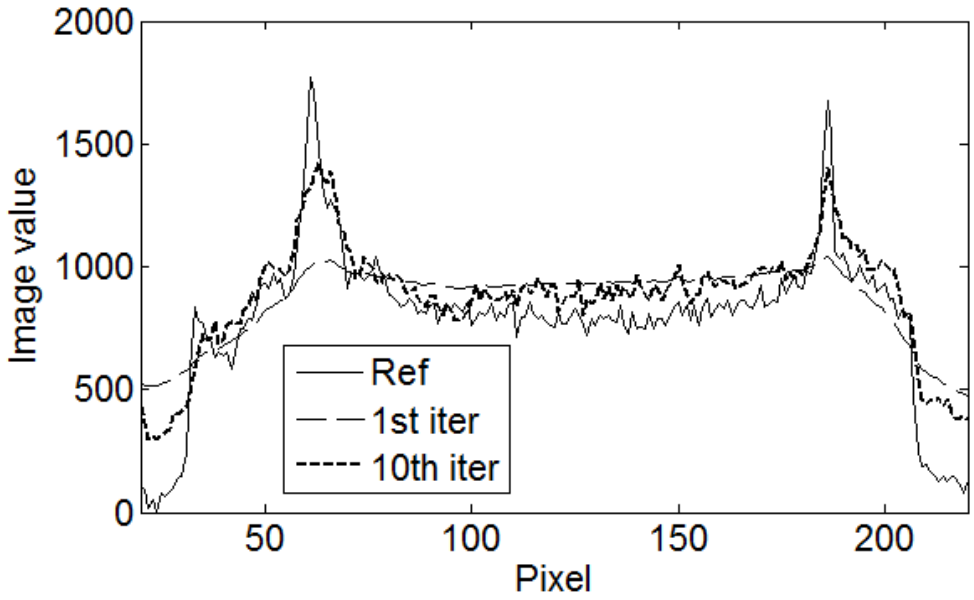


Figure 5.8: Image row plots for 1/3 of data compared to full data set of FDK reconstruction, at $f(x = 128, y, z = 128)$.

to a comparable degree. Continuing to later iteration may cause similar degradation of full data of this image as shown in Figure 5.2(c).

In addition, an area of the image containing some edges between bone and soft tissue has been selected to compare the reconstructed images and the reference FDK image. The selected area is marked in Figure 5.2(b). Figure 5.7 shows a plot of the norm difference between these areas (transaxial plane plot) with respect to iteration number for full and reduced data sets. The value of λ was set to 0.0146 for all reconstructions. It can be observed that, as expected, the rate of convergence increases when more projection data is available. However, for the full data set, using this value for λ, the reconstruction does diverge after around 13 iterations. For a half data set, the image quality as measured by this method does come close to that provided by the full data set, but a greater number of iterations are required.

Norm differences in projection data and norms of differences in tomographic images are generated as an indiction to be used in selecting the optimal number of iterations and relaxation parameters. In an ideal situation both these norms will be minimised.

Figure 5.9 is the plot of image mismatch errors against projection mismatch errors

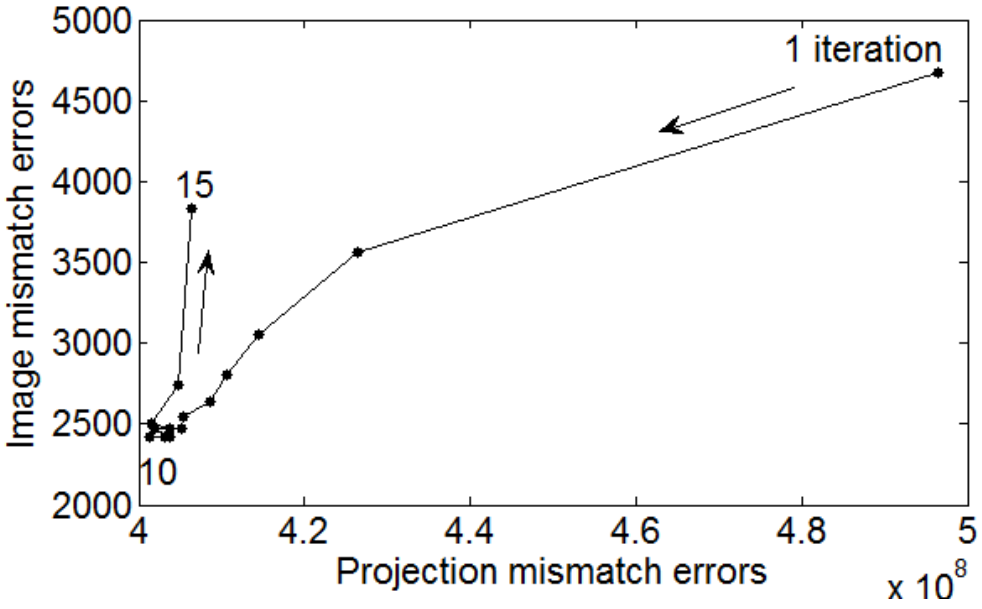


Figure 5.9: Image mismatch errors against projection mismatch errors for different iteration.

for different number of iterations for $\lambda = 0.0146$. As is shown, image errors and projection errors both decrease for the initial iteration. A good selection of number of iterations can be obtained when the plot achieves its minimum point.

When the number of iterations is decided, comparison of different relaxation parameters for full data are shown in Figure 5.10 for the same number of iterations obtained in Figure 5.9. The graph shows that, for a given range of relaxation parameters, the algorithm will converge quicker for a larger λ and becomes slower for a smaller λ. Larger λ requires fewer iterations and smaller λ requests more. However, different selections will not lead to the same output, where optimised settings for relaxation pa-

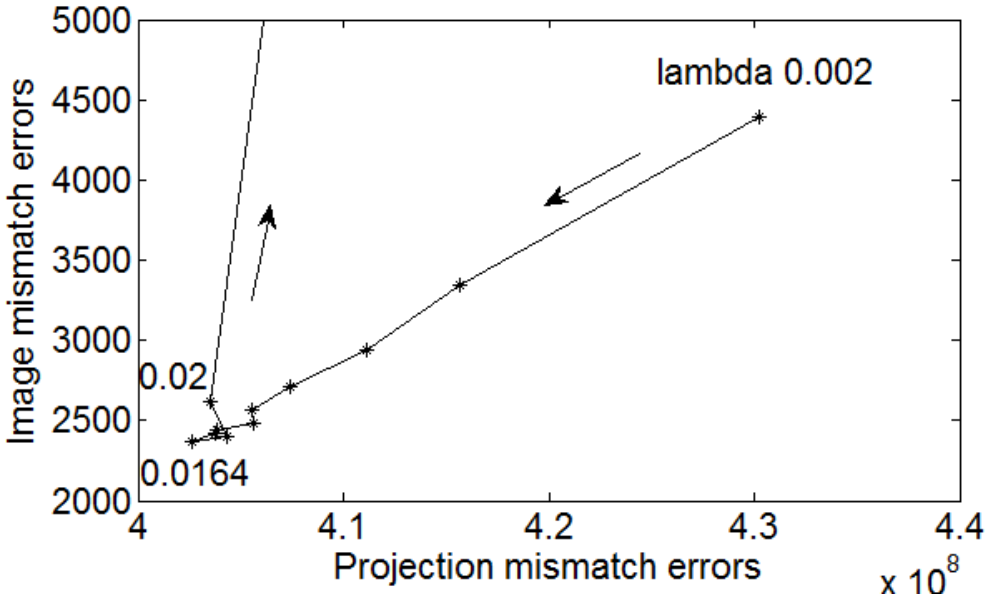


Figure 5.10: Relaxation comparison for the same number of iteration.

rameter and number of iterations exist. 3D plots of root mean square errors of image with respect to λ and the number of iterations as well as projection mismatch erros against λ and the number of iterations are then generated as shown in Figure 5.11. In an ideal case, best optimisation can be decided when both ΔE and ΔP are minimised, but as shown in the figures, when ΔE is minimal, ΔP may not be the minimised point and vice versa. In this case, ΔP, which presents the image quality, should be taken into account firstly rather than ΔE. The best selection may set as the following: while ΔP is minimal, ΔE falls into the domain R, where R is an acceptable range of projection errors. Or both ΔP and ΔE falls into the domain R_1 and R_2, where R_1 and R_2 are ac-

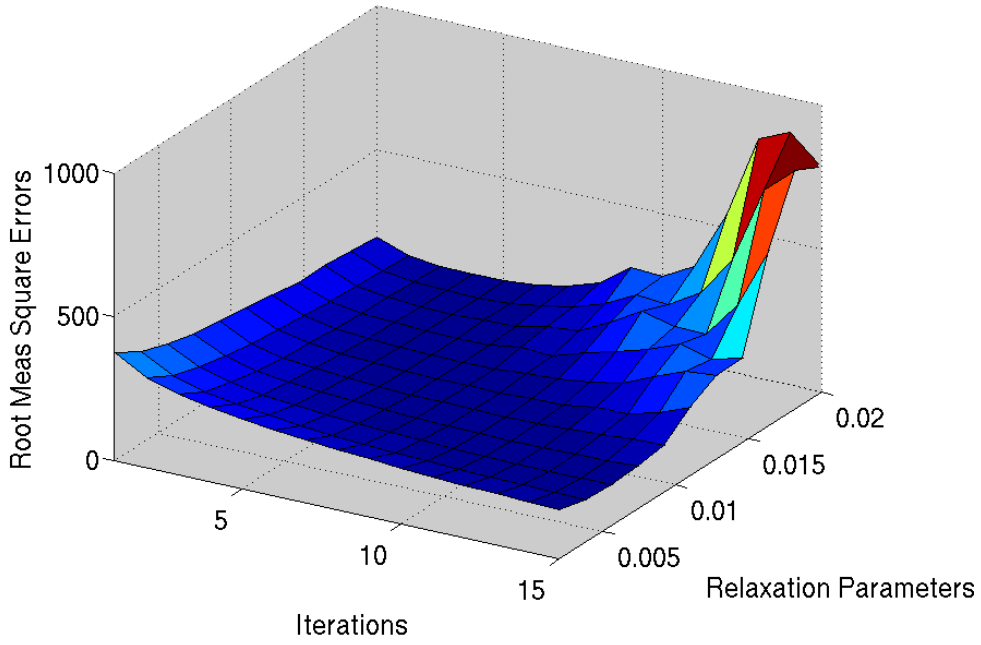


Figure 5.11: 3D plot of image errors against λ and i.

ceptable ranges of image and projection errors. The best selection is when $R_1 + R_2$ is minimal.

In addition, we have attached a table of the signal-to-noise ratio (SNR) of a region in the iterative method and FDK reconstructed images (region shown in Figure 5.1). Here, the SNR was determined inside regions of homogeneous absorption within the reconstruction, which could be a good quantitative way to assess the images quality. Moreover, SNR of reconstructions could be used to compare the quality of reconstruction with different number of projections. Table 5.1 shows the SNRs for FDK, and for

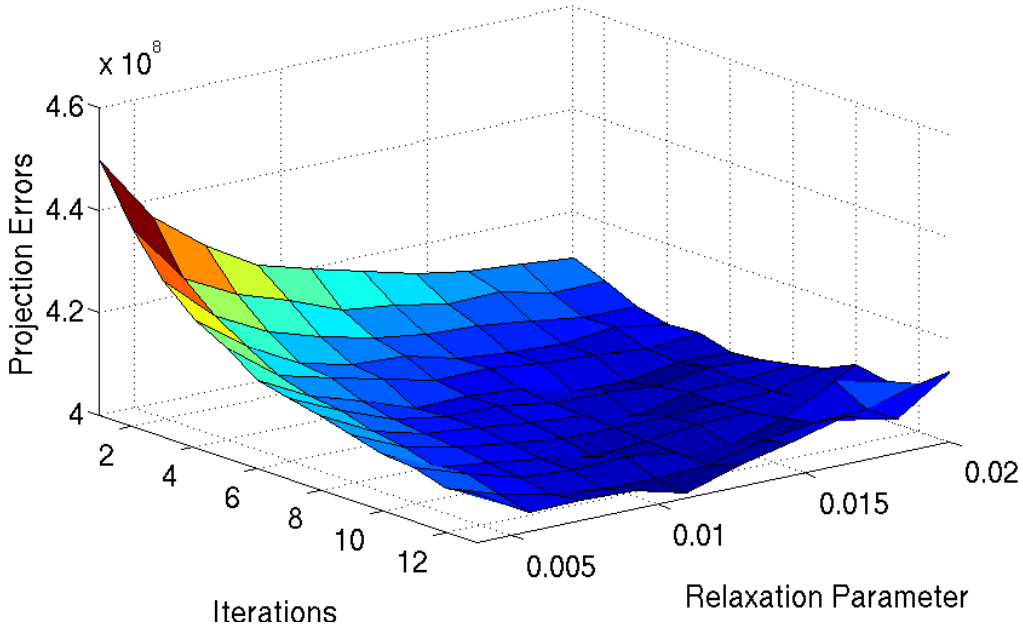


Figure 5.12: 3D plot of projection errors against λ and i.

iterative reconstructions with varying number of iterations and degree of down sampling, for some selected relaxation parameters. Although, instead of using the whole image (which usually is impossible to compare), only a region is selected for comparison, we can still capture some features of the reconstructions of the iterative method. However, this measure alone could not be considered to entirely capture the quality of the reconstruction. As indicated in the table, the SNRs for the MIDAS based iterative method with full data increase along iterations and decrease in later ones which is expected based on the discussion above. As the comparison are based on the optimised

settings for the MIDAS based iterative method with full data, the SNRs results for other reconstructions with reduced data vary. Some perform better (like 1/3 and 1/5) while some perform worse than expected (like 1/2 and 1/4).

Table 5.1: SNRs for different reconstruction settings (for the MIDAS based iterative method, $\lambda = 0.0146$) for region at $f(x = 113 : 136, y = 72 : 110, z = 128)$ in dB.

	2	4	6	8	10	12	14	16	18	20
FDK	15.98									
MIDAS(full)	14.34	15.79	16.57	16.83	16.62	16.16	14.78			
MIDAS(1/2)	14.59	14.62	14.85	15.00	15.40	15.71	15.97	16.10	16.10	16.10
MIDAS(1/3)	16.46	16.44	16.75	17.00	17.26	17.34	17.29	17.31	17.32	17.34
MIDAS(1/4)	14.82	14.70	14.84	15.71	15.43	15.52	15.64	15.70	15.73	15.67
MIDAS(1/5)	15.70	16.28	16.20	16.36	16.45	16.60	16.70	16.69	16.64	16.63

In order to strengthen the conclusion for the optimum value for λ while for the testing of 3D data reconstruction, a repeat of the experiment for a different set of clinical data in 2D is used. The convergence plot for this set of data for different selections of λ is shown in Figure 5.13. It can be observed from the figure that the selection of $\lambda = 0.0146$ is one of the best selections, the same as the conclusion obtained from the results shown in this chapter earlier.

Finally, we apply our results to clinical data as shown in Figure 5.14.

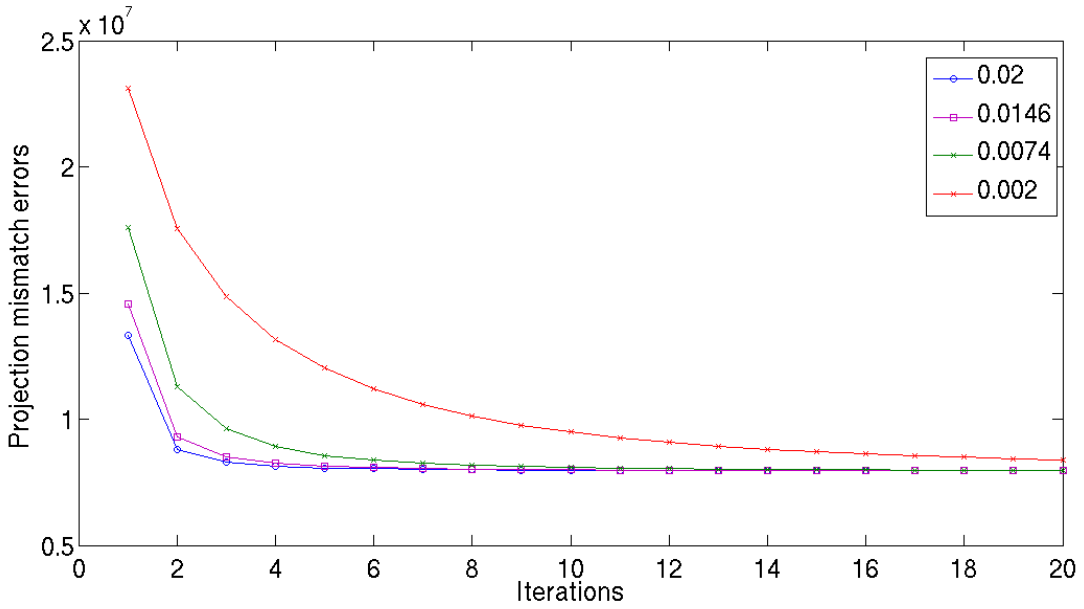
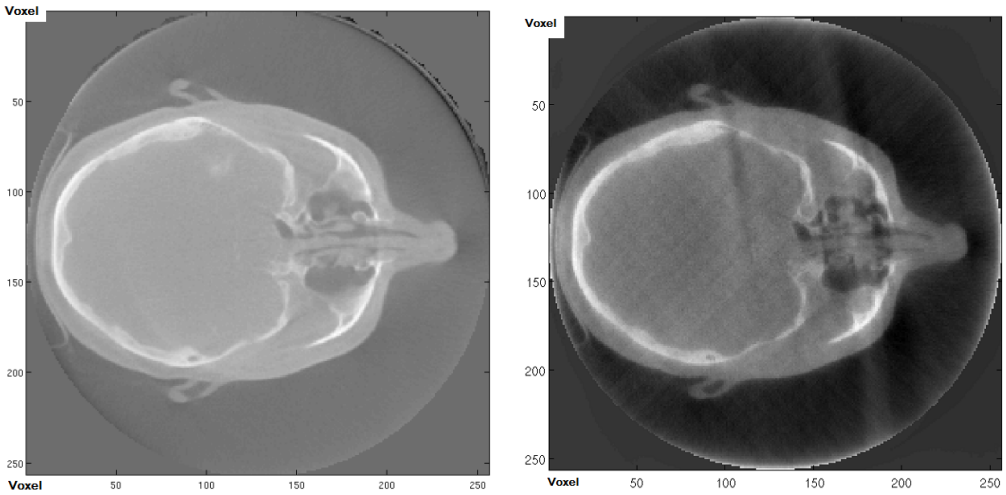


Figure 5.13: Convergence for λ of 0.02, 0.0146, 0.0074 and 0.00.



(a) FDK reconstruction (b) Results applied to clinical data.

Figure 5.14: Reconstruction using clinical patient data.

5.2 Conclusion

Iterative reconstruction using MIDAS based iterative method is investigated and implemented in MIDAS. Optimised results are presented. The settings of relaxation parameter and number of iterations are optimised based on reconstruction stability, convergence rate and image quality. The 1D image row plot shows the quality of the reconstructed image and contrast of edges. The transverse sum plot focuses on reconstruction for a specific area in comparison to the reference image. With these evaluations, a proper relaxation parameter and number of iterations can be determined and optimised, according to different situations and different amounts of available data. The results indicate that an optimised set of parameters can be found that both image errors and projection errors are minimal. With the results, 1/2, 2/3, 3/4 and 4/5 of projection data are removed symmetrically, and results suggest that with optimised parameters a reconstructed image suitable for clinical use can be obtained from reduced data sets.

Chapter 6

Comparative analysis of iterative methods

In previous chapter, MIDAS based iterative methods was introduced. Generally speaking, those iterative methods are not proper one compared with the classic iterative algorithms. Hence, standard linear algebraic iterative algorithms have been investigated in this chapter. We have used different iterative algorithms (ART, SART and OS-SART), focussing on convergence properties and image quality parameters (uniformity and noise). In addition, by adding more challenges to the work and for a fair comparison,

the iterative images were reconstructed without any denoising process. Further more, they were reconstructed with limited projection data (1/4) while the FBP image was reconstructed using full dataset. In particular we used a relaxation schedule which guarantees convergence.

6.1 Results and discussion

Results will be shown firstly on simulated data, then a more thorough study on experimental data, for 2D reconstructions. Then a cone beam 3D CT reconstruction result is attached.

6.1.1 Experimental setup

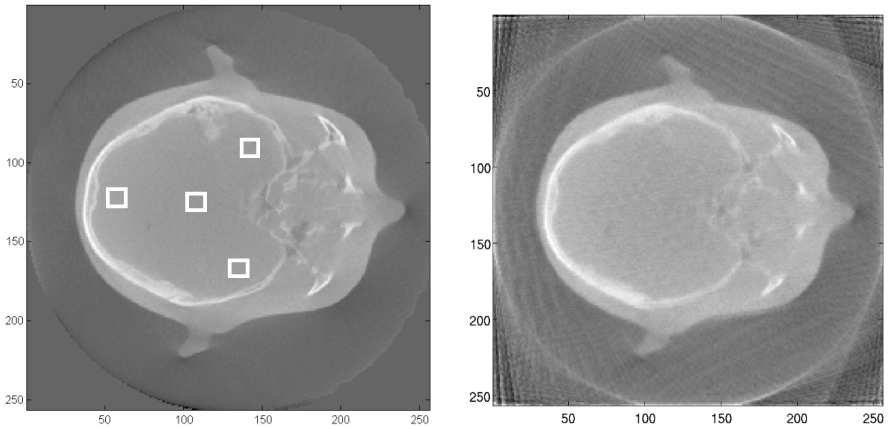
The experimental measured projection data were provided by North Western Medical Physics at The Christie Hospital in Manchester, which has an image guided radiotherapy research facility equipped with an Elekta Synergy® linac-integrated X-ray cone-beam CT system[1]. This consists of a kilovoltage X-ray tube and flat panel imager mounted onto a radiotherapy linear accelerator gantry at 90 degrees to the megavolt-

[1]Elekta Oncology, Crawley, UK.

age treatment beam (see Fig.1.1). A 'RANDO' anthropomorphic head phantom[1] was scanned to produce 360 X-ray projection images, approximately evenly spaced over an angular range of -100 to +100 degrees. Images were acquired at 100kV, 10mA and 10ms per projection, with total imaging dose of approximately 1.5mGy. Each projection image contains 512×512 pixels of dimension 1×1mm (note this is down-sampled from the system maximum resolution of 1024×1024). A 'reference image' was reconstructed using the COBRA cone beam software developers package from EXXIM[2]. This contains an implementation of FDK FBP, which is a useful benchmark for the iterative algorithms described in this chapter. In [42], image reconstruction for full and down-sampled projection data were studied. In this work, all reconstructions were implemented using 1/4 of the available projection data, simply to reduce computational cost and to make the reconstruction task more challenging. Fig.6.1(a) details reconstructions of the 'Rando' phantom using commercial clinical FDK software from 3D inversion and the iterative method (ART) applied from 2D inversion in this study. Note that additional filtering has not been implemented for the iterative reconstruction.

The computational time for each OS-SART iteration for a 3D reconstruction is

[1]The Phantom Laboratory, Salem, NY, USA.
[2]Exxim Computing Corporation, Pleasanton, CA, USA.



(a) 'Rando' phantom FDK reconstruction with full projection data from 3D inversion

(b) 'Rando' phantom ART reconstruction with 1/4 projection data from 2D inversion

Figure 6.1: Examples of experimental images.

shown in table 6.1. The computer used is 64 bit 3.33GHz Linux with ram of 32GB.

Table 6.1: Time cost for one OS-SART iteration with different block sizes

512 projections	256 projections	128 projections	64 projections	32 projections
3600s	7200s	14400s	28800s	57600s

6.1.2 Simulated data

In this section, an initial result is shown for reconstruction using ART, SART and OS-SART with a fixed relaxation schedule. Note that simulated data is used such that the problem is consistent.

The simulated phantom was generated in 2D with simple objects as shown in Fig.6.2. Fig.6.3 shows the plot of root mean square (RMS) reconstruction error against

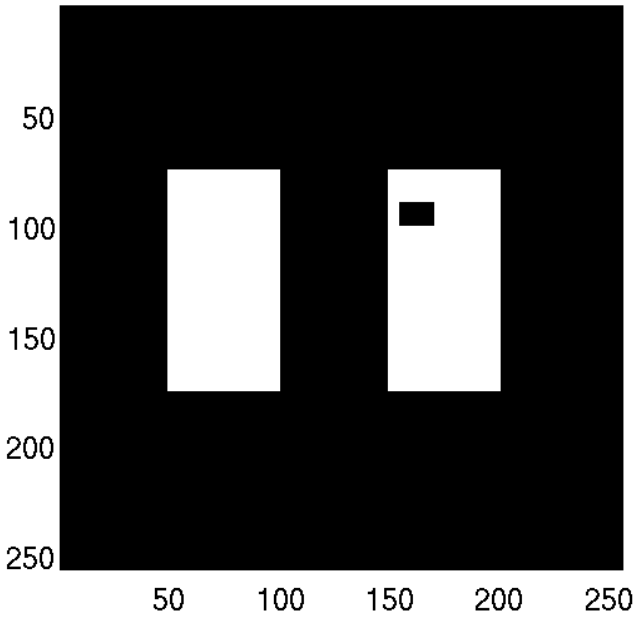
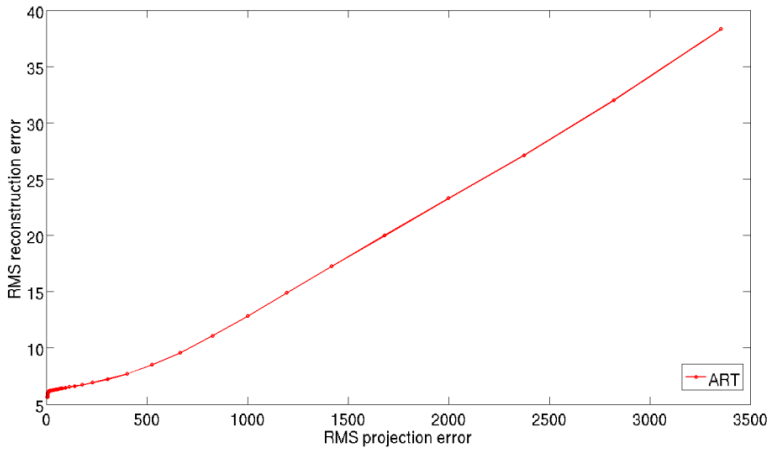
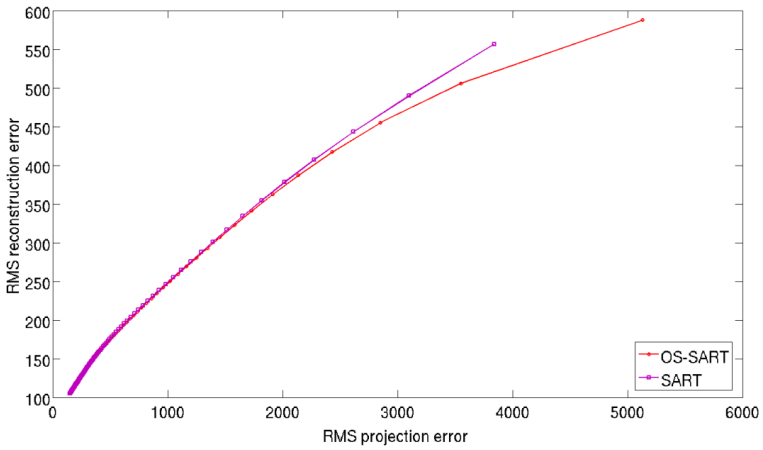


Figure 6.2: Simulated phantom.

root mean square (RMS) projection error for ART, SART and OS-SART. Because of different scales, they are shown in two plots. Points along the lines in the graphs denote errors after each cycle. RMS reconstruction errors were defined relative to the truth and λ was set to $\lambda = \sigma$, $\sigma = 0.1$. The purpose of this simulated study is to verify the monotonic convergence in root mean square reconstruction error for iterative algorithms with the fixed relaxation schedule for the consistent case. This monotonicity can be observed from the plots for ART, SART and OS-SART. As SART and OS-SART have potential for parallelisation compared to ART, and SART is a special case of OS-SART, a more thorough investigation of OS-SART is presented next.



(a) ART



(b) SART and OS-SART

Figure 6.3: RMS reconstruction error against RMS projection error for different algorithms.

6.1.3 OS-SART with experimental data

In this section, we investigate the behaviour of OS-SART. Firstly, we compare different relaxation schedules for a fixed relaxation parameter ($\alpha = 0$) but with different scaling factors σ. Convergence in RMS reconstruction error is guaranteed for a consistent problem providing $\sigma \in (0, 2)$. However note that the experimental data yields an inconsistent problem and convergence is not guaranteed. Following this, we compare the effect of a changing relaxation schedule $\alpha \in (0, 1]$. Finally, we present the study of OS-SART for different block sizes and the introduction of image quality measurements.

First, Fig.6.4 is the plot of RMS reconstruction error against RMS projection error for OS-SART with a traditional fixed relaxation schedule ($\alpha = 1$) with different scaling factors σ (with one full projection for each block). As the fixed relaxation schedule has been well studied, following is a brief summary of the main characteristics. Iterative reconstruction algorithms with a larger λ will quickly converge but often produce image degradation due to "semi-convergence". With a smaller λ, convergence will be slower but often eventually generate a smoother image. Reconstruction heavily de-

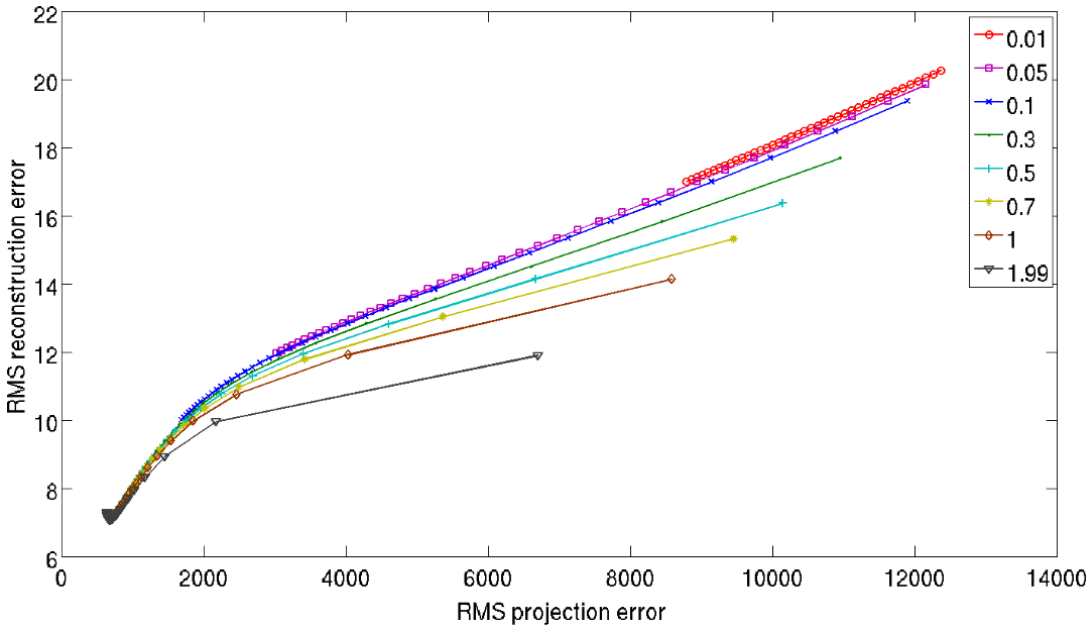


Figure 6.4: RMS reconstruction error v. RMS projection error for a fixed relaxation parameter with different scaling factors $\sigma \in \{0.01, 0.05, 0.1, 0.3, 0.7, 1, 1.99\}$.

pends on the correction part of the algorithm, requiring a combination of selection of relaxation schedule and number of cycles. The lower the λ the more cycles are needed, and vice versa. Obviously, the semi-convergence property can also be observed from the plot when $\sigma = 1.99$, where semi-convergence is defined as an increase in the RMS reconstruction error with successive cycles.

For the fixed relaxation parameter, convergence is not guaranteed as defined in Sec.3.2.1.2. However, a relaxation parameter schedule with λ changing with each cycle can guarantee convergence as shown in Fig.6.5. In this figure, OS-SART (one

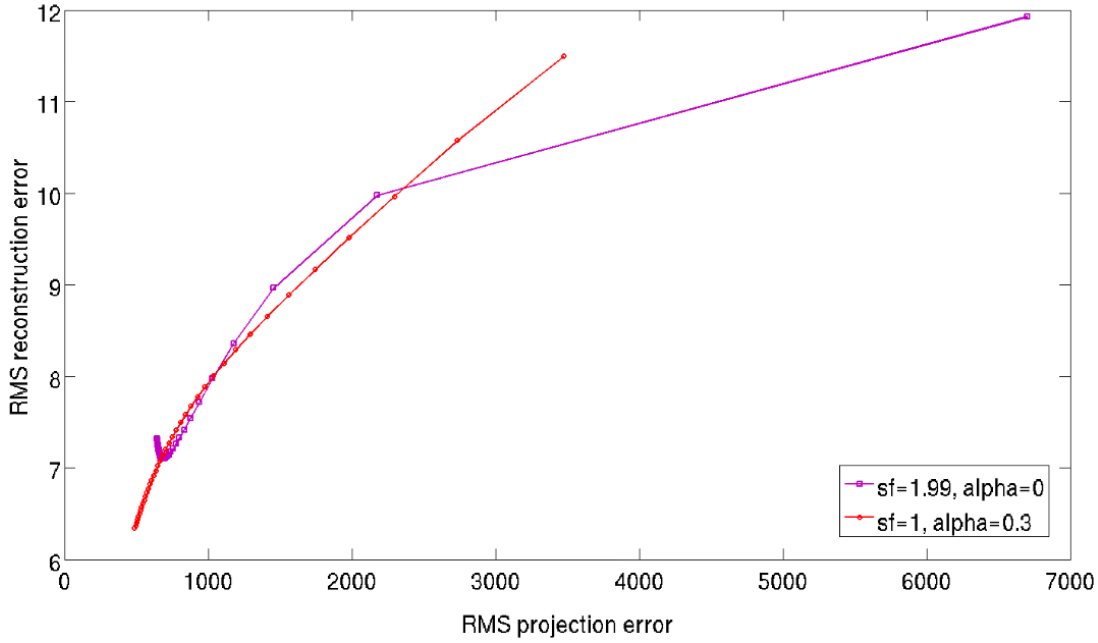


Figure 6.5: RMS reconstruction error v. RMS projection error for a fixed relaxation schedule against changing relaxation schedule. (sf is σ and alpha is α)

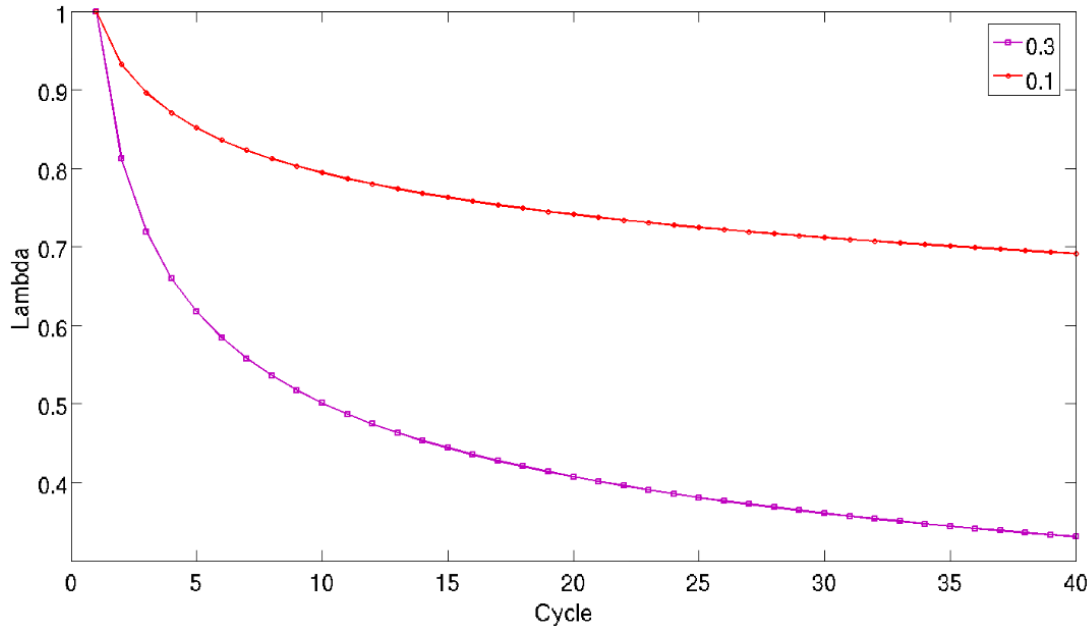


Figure 6.6: Different relaxation schedules for λ with $\alpha \in \{0.1, 0.3\}$.

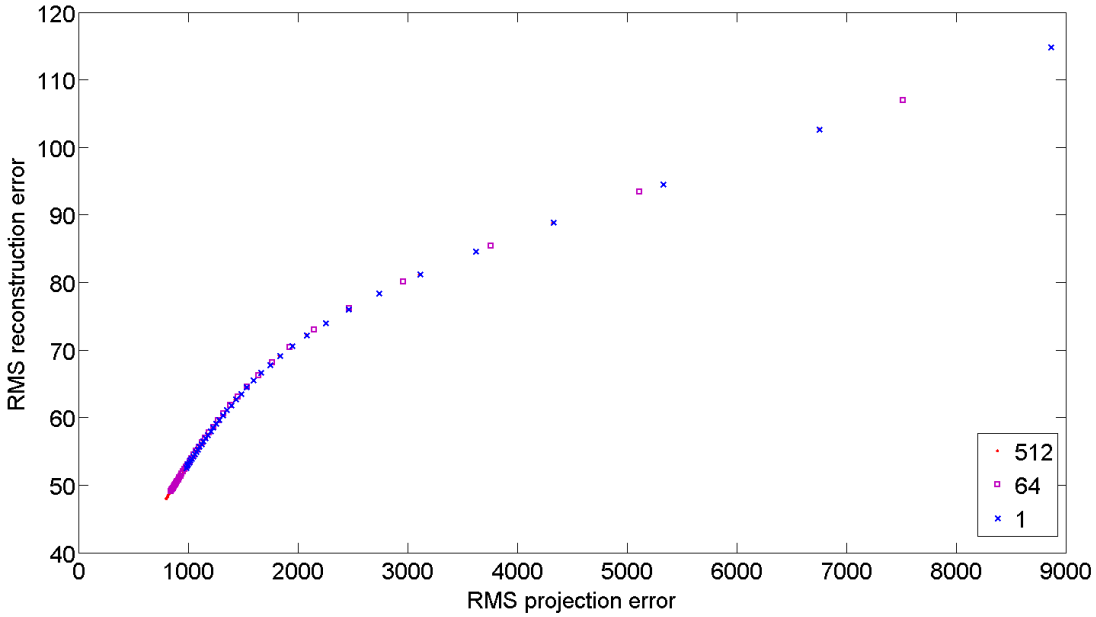


Figure 6.7: RMS reconstruction error v. RMS projection error for different block sizes with $\sigma = 1$ and $\alpha = 0.3$.

projection per block) with $\alpha = 0.3$ was implemented. The plot of the fixed relaxation parameter schedule diverges in later cycles while for the changing relaxation schedule, both the RMS and projection errors are monotonically decreasing.

Next the reconstructions were tested for different schedule suing the same scaling factor $\sigma = 1$ but with different $\alpha \in \{0.1, 0.3\}$. The schedules are as shown in Fig.6.6. With higher α, a faster change in λ was observed.

Different block sizes for OS-SART were investigated at the same time. Block sizes ranging from 512 rays (a full projection) to 1 ray were tested. Fig.6.7 is the plot

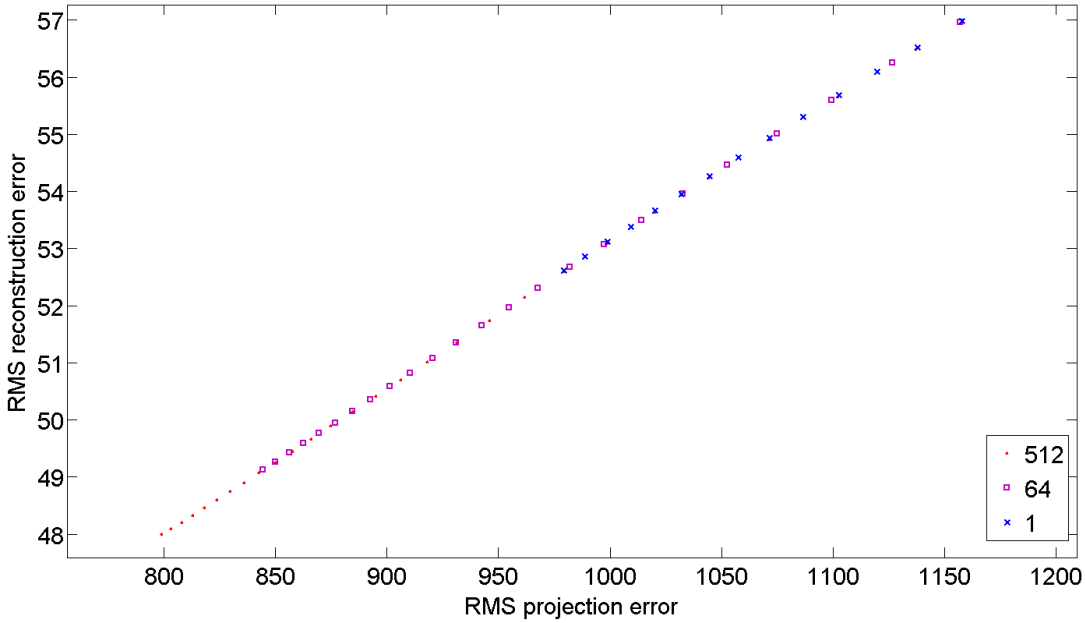


Figure 6.8: Zoom view of Fig.6.7.

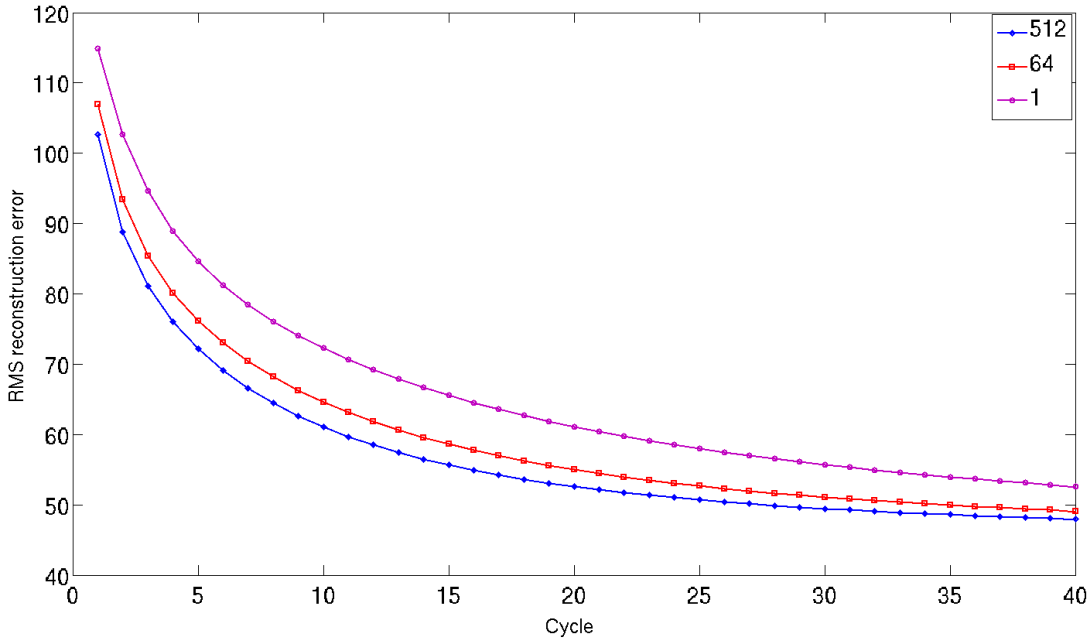


Figure 6.9: RMS reconstruction error v. cycle index for different block sizes with $\sigma = 1$ and $\alpha = 0.1$.

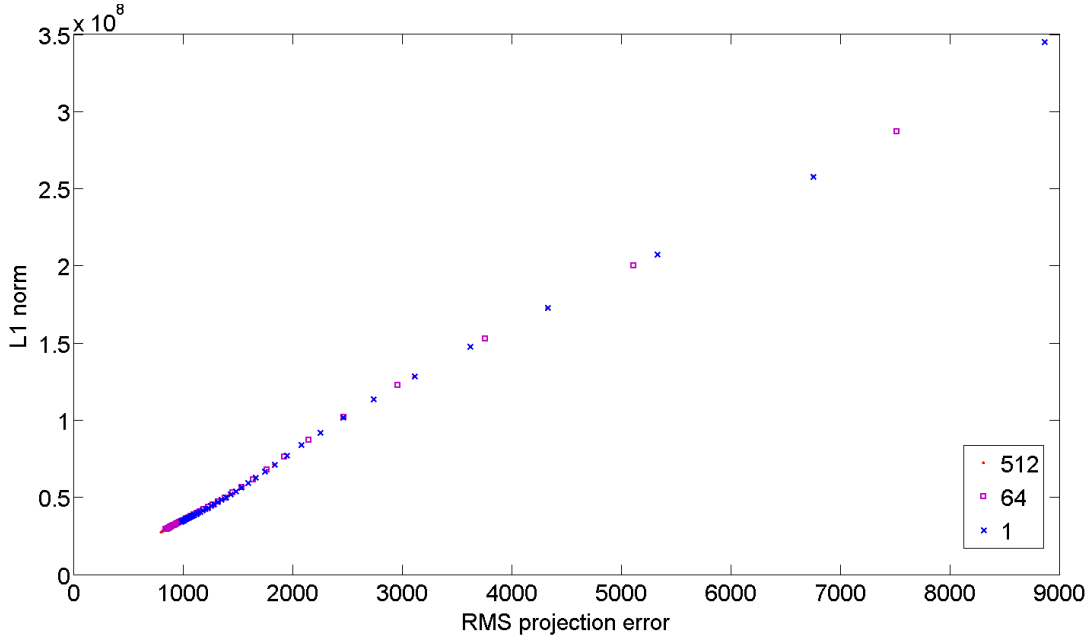


Figure 6.10: L1 norm v. RMS projection error for different block sizes with $\sigma = 1$ and $\alpha = 0.3$.

of RMS reconstruction error against RMS projection error with different block sizes for OS-SART. For these experiments, σ was set to 1, α was 0.3 and the number of cycles was 40. To show the effect of block size in convergence of the algorithms, we have zoomed the image of Fig.6.7 in Fig.6.8, which suggests faster convergence for a larger block sizes. Fig.6.9 shows the rate of convergence for different block sizes under the same parameters except $\alpha = 0.1$. Broadly, the reconstruction with block-size of 512 rays converges faster than the one with smaller block-sizes. Fig.6.10 and Fig.6.11 are the L1 and L2 norm of the reconstruction against RMS projection errors

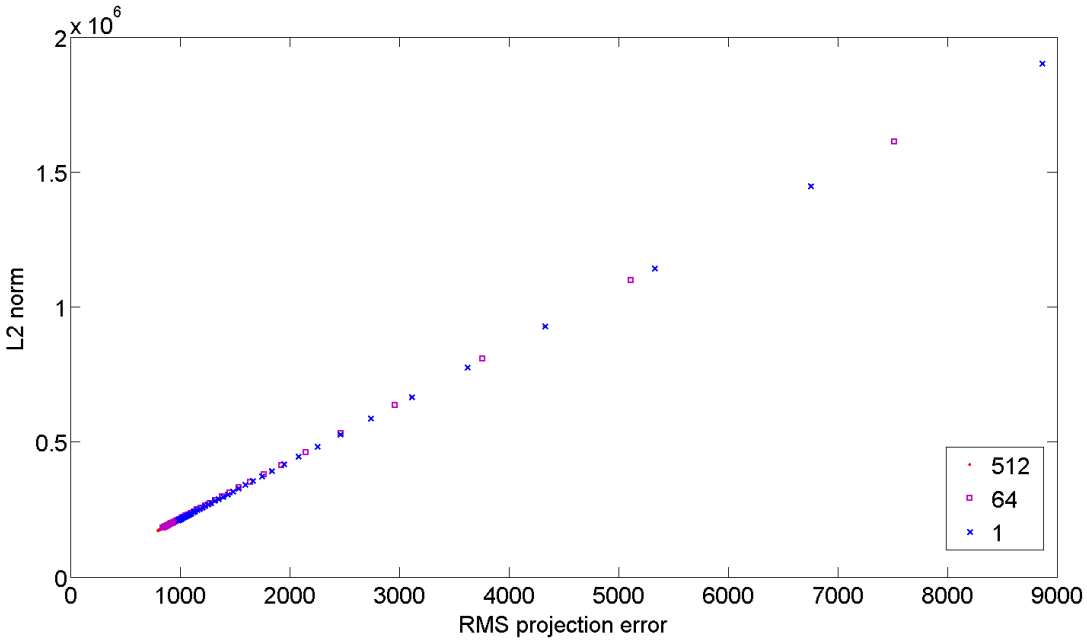


Figure 6.11: L2 norm v. RMS projection error for different block sizes with $\sigma = 1$ and $\alpha = 0.3$.

respectively. Similar to Fig.6.9, observations suggest that for the same number of iterations, reconstruction with larger block sizes has a lower RMS reconstruction error and RMS projection error.

6.1.4 Using image quality parameters

The tests were then repeated using image quality assessment, where uniformity and noise parameters were used. Fig.6.12 and Fig.6.13 are the plots of reconstructed image quality against RMS projection error with different block sizes for OS-SART. Image

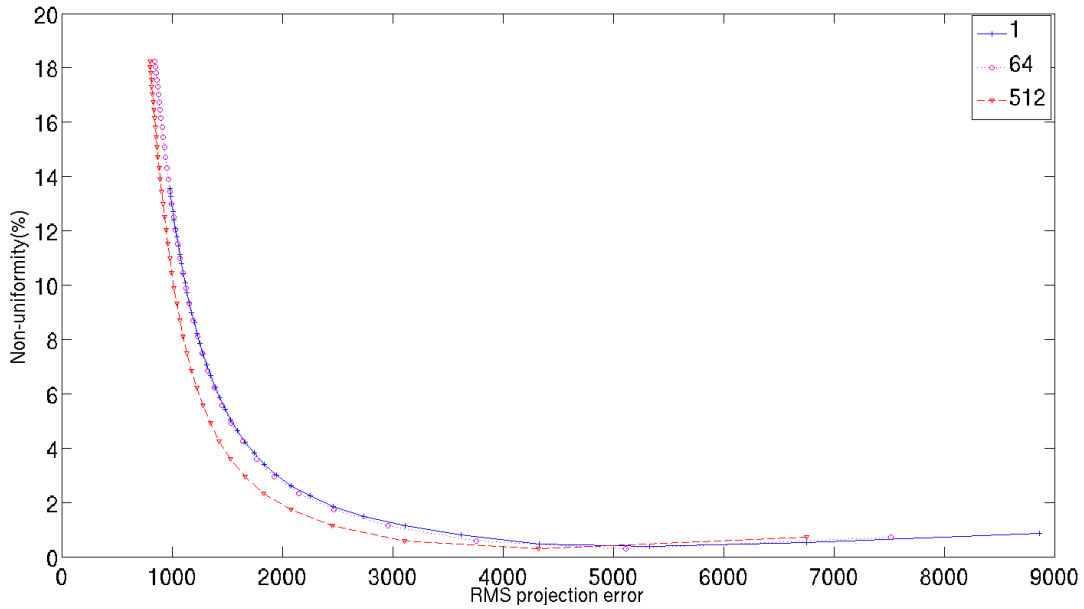


Figure 6.12: Non-uniformity function plotted v. RMS projection error for block sizes $\in \{1, 64, 512\}$ with $\sigma = 1$ and $\alpha = 0.3$.

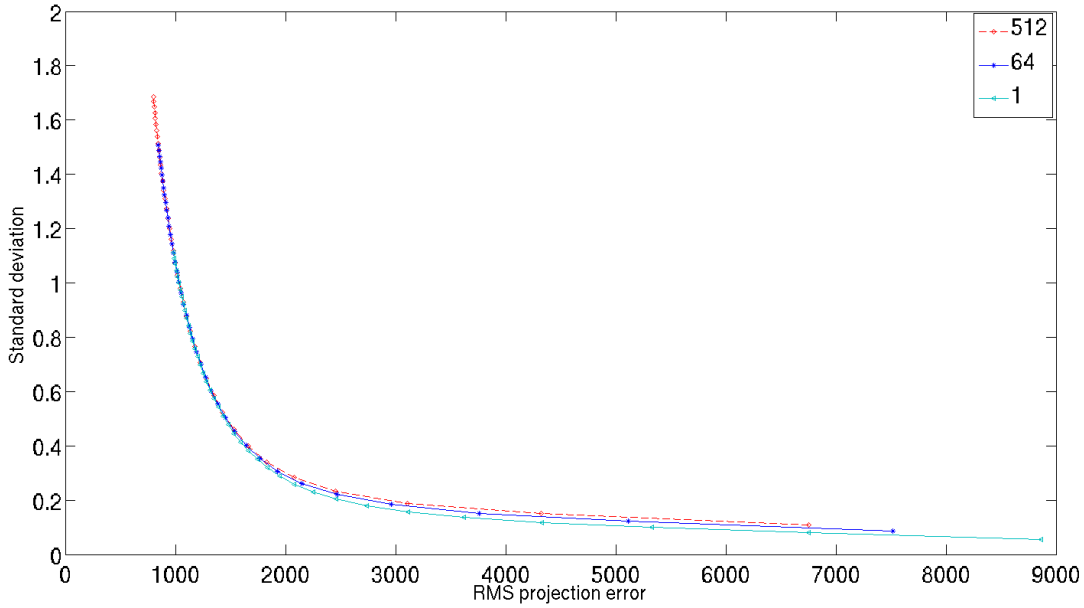


Figure 6.13: Noise function plotted v. RMS projection error for block sizes $\in \{1, 64, 512\}$ with $\sigma = 1$ and $\alpha = 0.3$.

quality parameters are plotted against RMS projection error rather than cycle index in an attempt to "normalise" for different rates of convergence in λ. As can be observed, after the initial few cycles, large block sizes give greater uniformity for given RMS projection error, but also greater noise. Fig.6.14 combines RMS reconstruction error and image noise according to Eq.4.9 with different μ. In addition, a surface plot is shown in Figure 6.15. A best selection of RMS reconstruction error and image noise can be obtained to have the minimal $h(\boldsymbol{x};\mu)$. However, the choice of μ is subjective and should be optimised depending on the specific context, e.g. to maintain a tolerance on noise or uniformity level, or for limited angle geometry. As explained in Sec.4.6, there is a stationary point as shown in the figure.

6.1.5 Fully 3D reconstruction

The experimental results obtained from the previous section were next applied to a set of 2D projections of CBCT data. A 3D reconstruction of 256×256×256 voxels with resolution 1mm in each direction was produced using OS-SART (one projection per block, $\sigma = 1, \alpha = 0.1$). An axial slice is presented in Fig.6.16(b). The 3D forward projection was implemented once and the projection matrices were stored and loaded

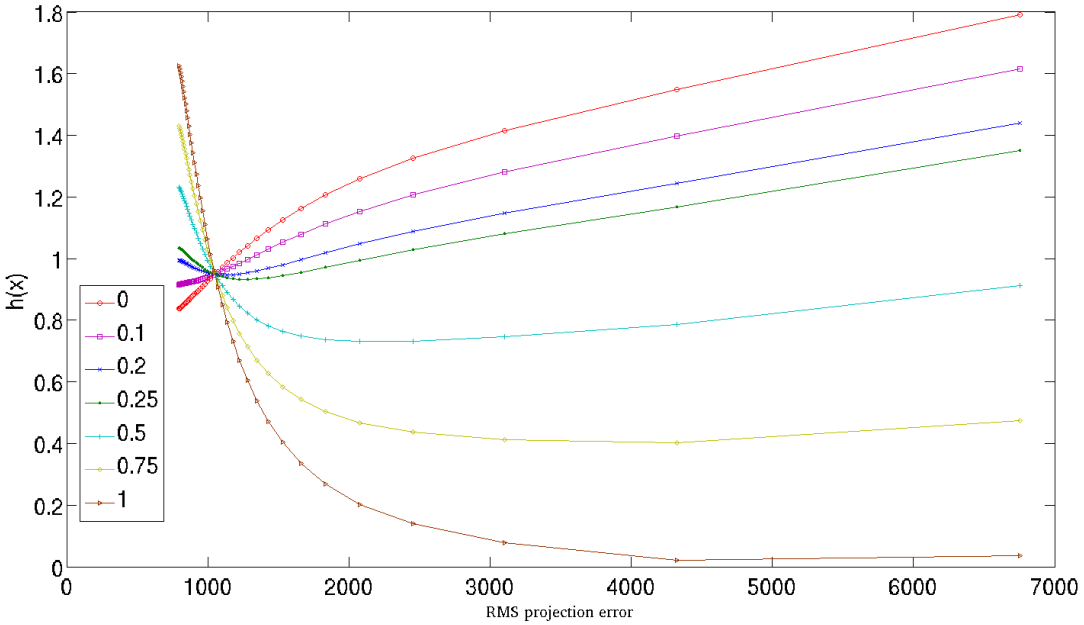


Figure 6.14: $h(\boldsymbol{x};\mu)$ (Eq.4.9) plotted against RMS projection error with different $\mu \in \{0, 0.1, 0.2, 0.25, 0.5, 0.75, 1\}$.

as necessary. This reduces the computational load but increases the disk storage and memory requirements. In future work, we plan to mitigate the computational challenges using techniques from high performance computing. Since these projection matrices are extremely sparse, especially in 3D, we expect gains can be made through more advanced numerical methods, e.g. preconditioned conjugate gradient methods [64].

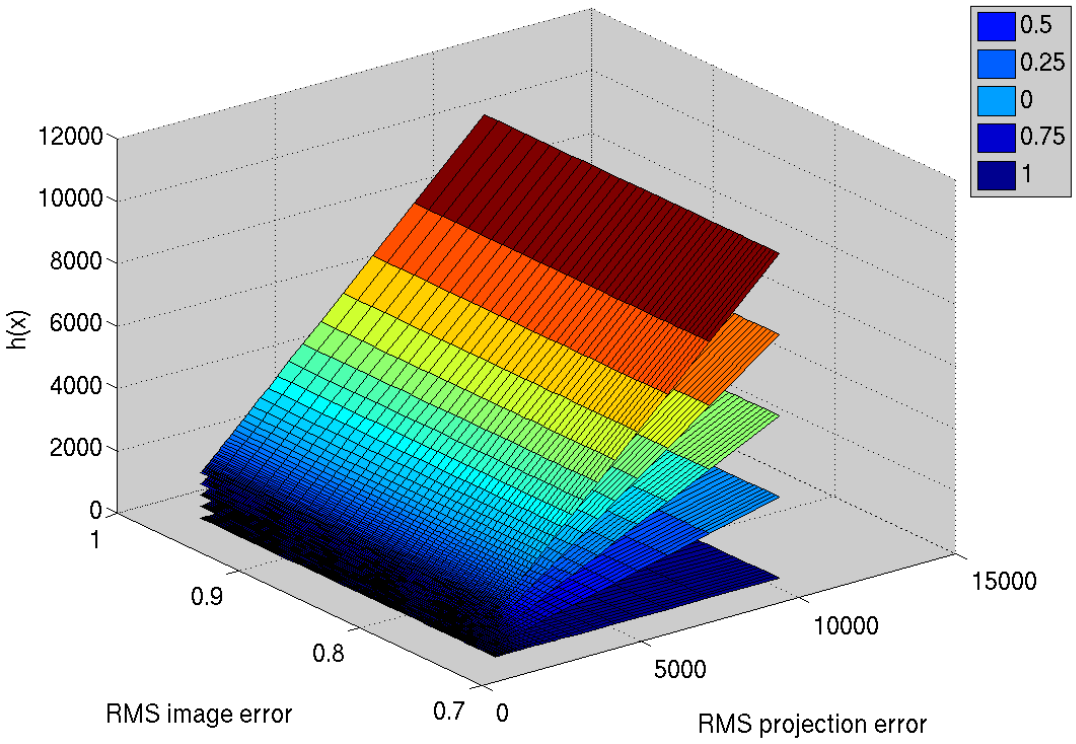
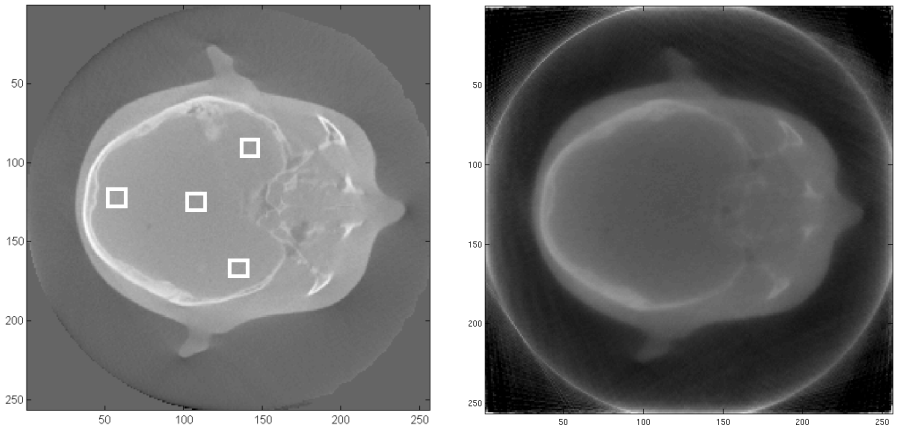


Figure 6.15: 3D plot of function $h(\boldsymbol{x};\mu)$ with different $\mu \in \{0, 0.25, 0.5, 0.75, 1\}$.



(a) 'Rando' phantom FDK reconstruction with full projection data from 3D inversion

(b) 3D CBCT reconstruction with axial slice $z = 128$ using OS-SART with $\sigma = 1, \alpha = 0.1$.

Figure 6.16: Examples of reconstructed images.

6.2 Conclusions

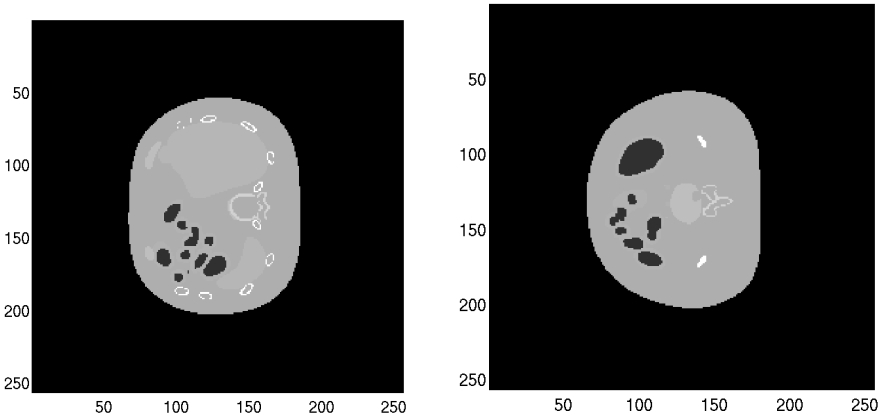
Linear algebraic iterative algorithms have been well studied. However, there is a tendency to study them theoretically or practically. Aiming to bridge this gap, a practical study was conducted based on theoretical results in convergence. In this chapter, we have used different iterative algorithms (ART, SART and OS-SART), focussing on convergence properties and image quality parameters (uniformity and noise). By adding more challenges to the work and for a fair comparison, the iterative images were reconstructed without any denoising process. In addition, they were reconstructed with limited projection data (1/4) while the FBP image was using full dataset. In particular we used a relaxation schedule which guarantees convergence. 2D reconstruction results were presented on simulated and experimental data, and a 3D reconstruction using CBCT data. As SART has potential for parallelisation compared to ART, while SART is a special case of OS-SART, OS-SART was investigated more thoroughly. From a convergence point of view, for OS-SART, the larger the block size, the faster the convergence. However image qulaity typically degrades as the algebraic algorithms converge. A trade-off between convergence and image quality should be made

to maximise the performance of iterative algorithms. A simple function was used to implement this trade-off.

Chapter 7

Block conjugate gradient type method

Standard iterative methods are well studied in CBCT reconstruction and as mentioned in previous chapters. However, for CBCT reconstruction, CG-type iterative methods are rarely investigated. The main challenge in CBCT reconstruction is that the matrix A is usually extremely large but very sparse. Limited or sparse data reconstruction are required. In addition, high performance computing may also be necessary to speed up the calculation. One of the advantages of CG-type iterative methods is that the convergence usually is fast. However, CG-type methods for CBCT reconstruction require storing the whole weighting matrix A before the actual calculation. A large amount of



(a) Visualised NCAT phantom at plane $f(x, y, z = 128)$.

(b) Visualised NCAT phantom at plane $f(x, y, z = 60)$.

Figure 7.1: True NCAT phantom visualisation.

computer memory is needed. This prevents the use of CG-type methods on commodity computers. In this chapter, one of the CG-type methods, LSQR was used for CBCT reconstruction. It demonstrates that LSQR can be used for CBCT reconstruction. Furthermore, block LSQR is proposed to overcome the memory capacity problem in 3D reconstruction. By using this proposed algorithm, CBCT reconstruction can immediately be used in high performance computing. And, the advantage of this algorithm will enable the use of personal commodity computers for CBCT reconstruction.

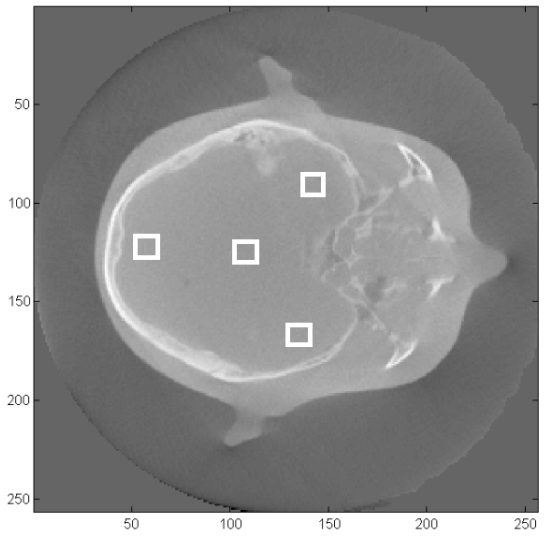
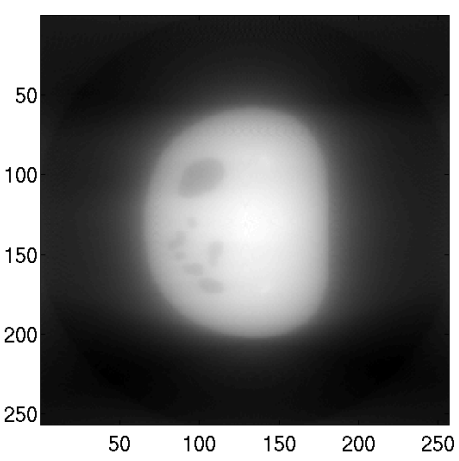


Figure 7.2: The head phantom reconstructed with FDK FBP.

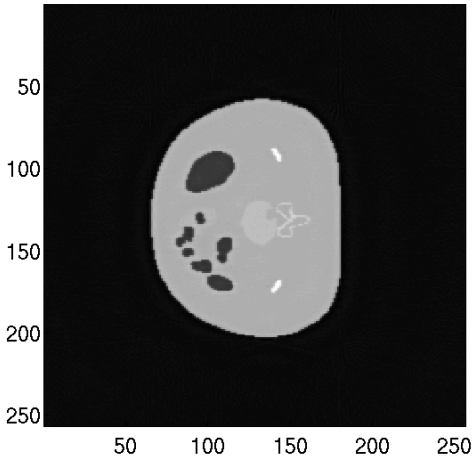
7.1 Full data set using LSQR

In this section, a standard LSQR algorithm was applied to full data set for reconstruction. Fig.7.1 shows two different slices of the true NCAT phantom for reconstruction testing and Fig.7.2 shows the head phantom reconstructed by using FDK filtered back projection.

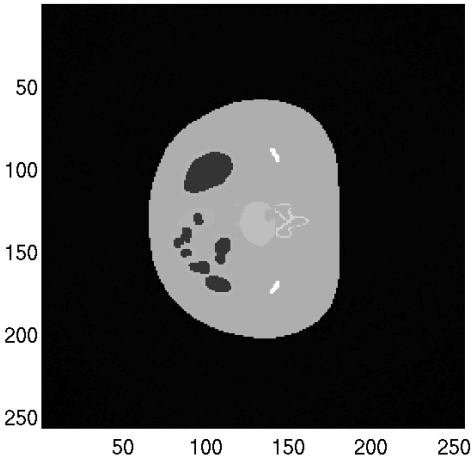
In Fig.7.3, there are four image reconstructions using LSQR at different iteration number and Fig.7.4 is the residual norm convergence plot of the reconstruction. This is a simulated phantom, so it would be considered as a noise free problem. As can be observed both from the reconstructed image and convergence plot, the more iter-



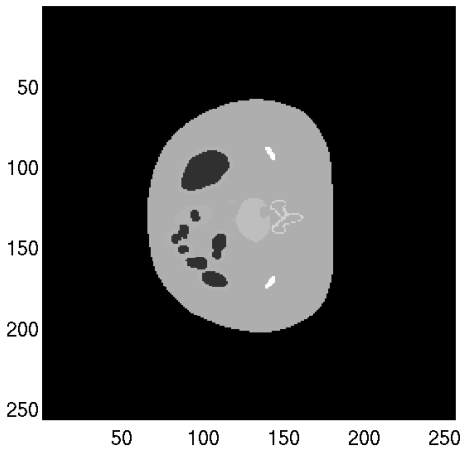
(a) Visualised NCAT phantom at plane $f(x, y, z = 60)$ (2 iterations).



(b) Visualised NCAT phantom at plane $f(x, y, z = 60)$ (15 iterations).



(c) Visualised NCAT phantom at plane $f(x, y, z = 60)$ (40 iterations).



(d) Visualised NCAT phantom at plane $f(x, y, z = 60)$ (100 iterations).

Figure 7.3: Full data set LSQR for NCAT phantom.

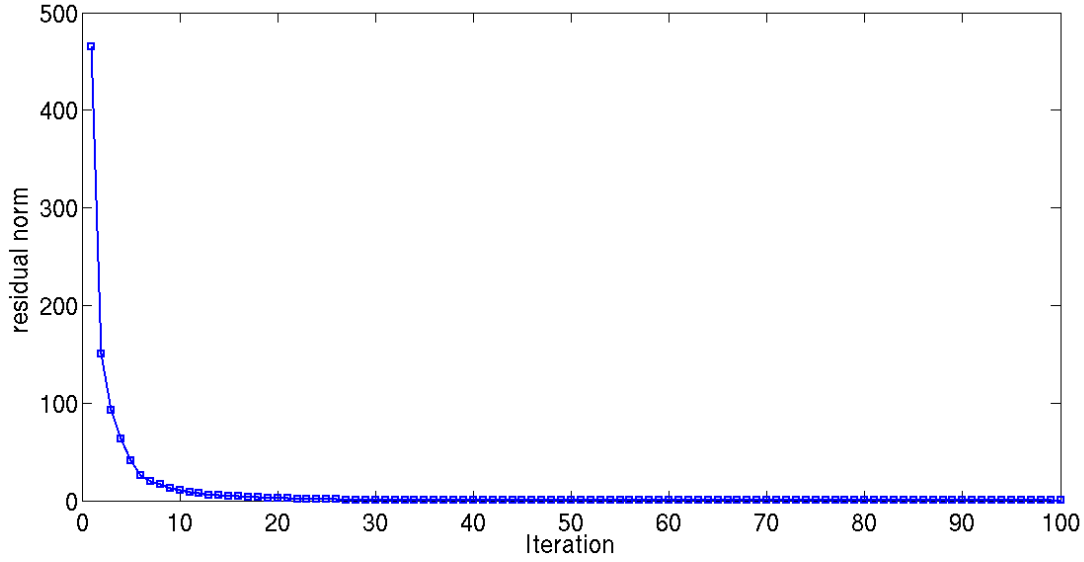


Figure 7.4: LSQR Convergence plot (100 iterations).

ation implemented, the clearer the reconstruction (differences between Fig.7.3(a) and Fig.7.3(b)). However, the image will not improve too much at later iteration. From the convergence plot, the residual norm reaches almost zero after around 20 iterations. Hence, there will not be too much differences between Fig.7.3(c) and Fig.7.3(d).

Fig.7.5 and Fig.7.6 are reconstruction examples by using standard LSQR with full data set. The reconstruction was not that good for the head phantom even full data set was used. This is because in this scenario the problem is not noise free leading to the inconsistent reconstruction. For the NCAT phantom, standard LSQR reconstruction using full data set reconstructs the image almost as the same quality as the true phantom

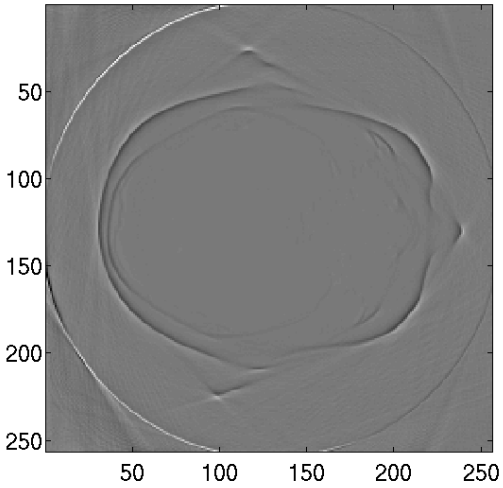
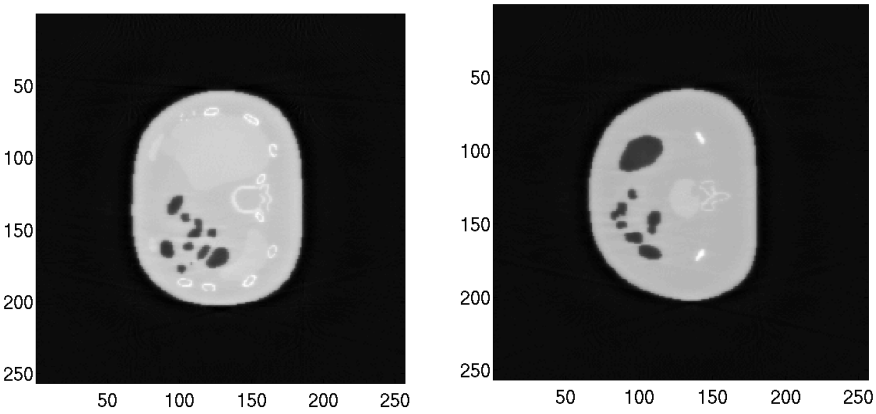


Figure 7.5: Full data set LSQR for brain phantom.



(a) Visualised NCAT phantom at plane $f(x, y, z = 128)$.

(b) Visualised NCAT phantom at plane $f(x, y, z = 60)$.

Figure 7.6: Full data set LSQR for NCAT phantom.

example.

7.2 Block LSQR

As mentioned previously, due to the large memory required for storing a 3D weighting matrix A, only 1/20 of a 360-full scan could be used under the constrain of the our current computer memory capacity (32GB). We have implemented sequential LSQR, in which projection matrices were used sequentially in LSQR algorithm and that did not proived a good convergence and the reconstrucion results were poor. A modified version of the LSQR is proposed which is the block LSQR. Block LSQR provides a solution for the constrain of the memory capacity problem. Even given a extremely large 3D weighting matrix A (e.g. $360 \times 512^2 \times 512^3$), block LSQR has the ability to handle it. In addition, block LSQR could run parallelly which can potentially increase the reconstruction speed dramatically.

In previous section, we discussed the memory capacity problem when handling large weighting matrix A. With the computer memory constrain, a computer with 32GB memory could only handle matrix A for 2D with the size of $360 \times 512 \times 512^2$

or 3D with the size of $360 \times 512^2 \times 512^3$. In the 3D case, it could only handle 1/20 of the full scan data, which limits its use in application. Block LSQR was then modified to overcome this issue. Essentially, block LSQR uses LSQR reconstruction but with a memory management capacity. If the data could be loaded at once (e.g. for 2D full scan or 1/20 of 3D full scan), block LSQR acts as the traditional LSQR. If the data must be loaded blockwise (e.g. for 3D full scan), block LSQR will firstly divided the data into blocks according to current computer memory capacity. With our computer capacity, a 3D full scan will be divided into 20 blocks. After that, it will load the divided data one by one for blockwise matrix-vector multiplications. In this way, the Algorithm 1 is the pseudo code for the implementation.

Algorithm 1 Block LSQR

```
Require: Scanning system parameters
  if Data can be loaded all at once then
    Execute LSQR
  else if Data must be loaded blockwise then
    if Parallel computing is possible then
      Perform matrix vector product parallelly
    else
      for 1:end do
        Load block data
        Perform matrix vector product
      end for
    end if
    Execute LSQR
  end if
```

The computational time for each iteration in LSQR are shown in table 7.1.

Table 7.1: Time cost for each iteration with LSQR and block LSQR

LSQR + loading data(once)	block LSQR (data loading on the fly)
50s+235s	1280s for each block

7.3 Results for proposed block LSQR

The proposed blockwise LSQR algorithm was used in 2D reconstruction with full data set for verification. In this following example, the matrix A has been divided into several sub-matrices. Fig.7.7 is the reconstruction and it demonstrates that the proposed algorithm is able to reconstruct the image correctly. The result is the same as the Fig.7.6(b) in which full matrix was used at once.

Now we consider the case for severely limited data set. Results are shown first for a 2D case NCAT [65] on simulated data, aiming to demonstrate that the more blocks used, the better the reconstruction. The results will then be presented for a 3D case on the same simulated data as with the 2D case. Finally, a more thorough study on the 3D 'Rando' head phantom is presented. In addition, image quality assessment (uniformity) is used to evaluate the reconstructed images. Note that, for a fair comparison,

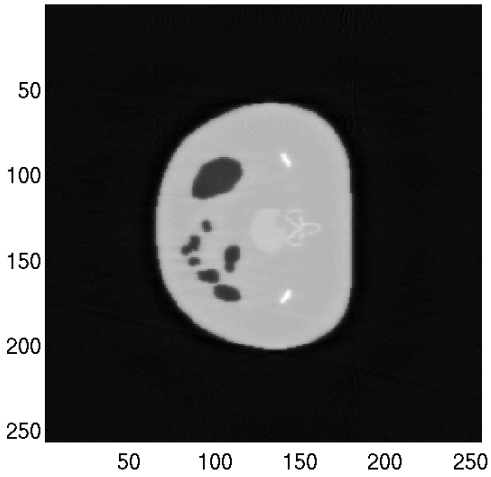


Figure 7.7: Full data set LSQR for NCAT phantom $f(x, y, z = 60)$ using blockwise method.

no filter was applied to the reconstructed images. Due to computational constraints, the reconstructed images look poor compared to the reference original images simply because only 1/10 (36 projection directions) or 1/20 (18 projection directions) of the available projection data were used in the 3D case.

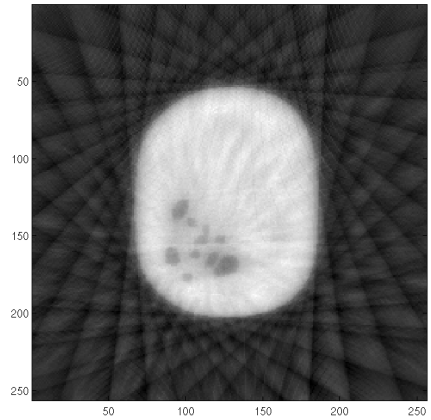
In the following, results are shown for the NCAT chest phantom [65] simulated data, for both 2D fan beam and 3D cone beam reconstruction. The aim is to demonstrate that the more blocks used, the better the reconstruction. Next, a more thorough study on the 3D 'Rando' head phantom is presented. In this case, image quality assessment (uniformity) is used to evaluate the reconstructed images. When no regularization is used, the reconstructed images look poor compared to the original reference

image. Results are therefore presented using Tikhonov regularization, with an empirically chosen regularization parameter $\alpha = 10^{-3}$ in (3.20). This significantly improves the quality of the reconstructed images.

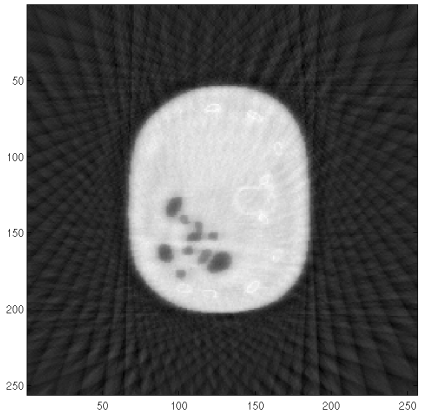
Even though this is usually not the most efficient choice for large scale computations, the algorithm has been implemented in the popular software package MATLAB. In all cases, the images have been produced by running 10 iterations of LSQR. (This choice is justified below.) For fair comparison, no additional filter has been applied to the reconstructed images.

7.3.1 Fan beam (2D) reconstruction of the chest phantom

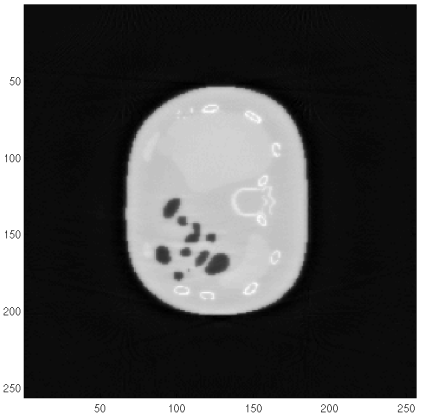
Images reconstructed from the 2D chest phantom data using blockwise LSQR are shown in Fig. 7.8. Clearly, the more data sets (scanned directions) used, the clearer the reconstructed image. Obviously, there is no advantage to using blockwise matrix-vector multiplications for 2D reconstruction, because the matrix A can be fully loaded into memory in this case. However, this approach becomes much more useful in 3D reconstruction, where memory requirements are well beyond the capacity of commodity computers. In this case, implementing the matrix-vector multiplications blockwise



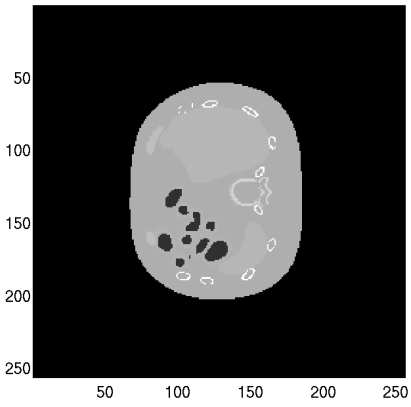
(a) 2D reconstruction using 1/20 data



(b) 2D reconstruction using 1/10 data

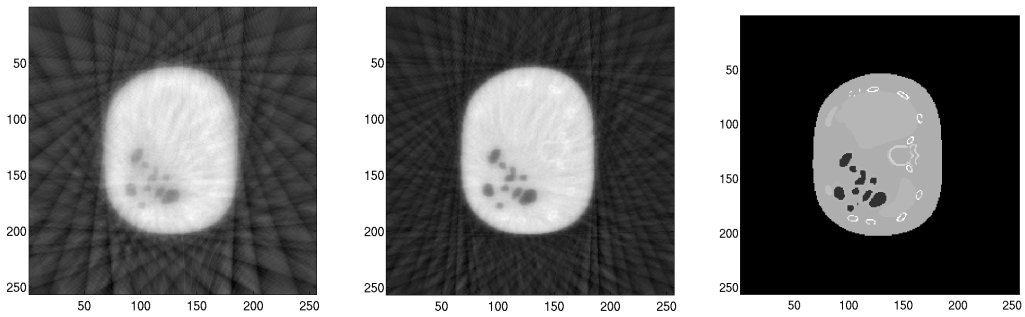


(c) 2D reconstruction using full data



(d) Chest phantom

Figure 7.8: Fan beam (2D) reconstruction of the chest phantom.



(a) 3D reconstruction using 1/20 data

(b) 3D reconstruction using 1/10 data

(c) Chest phantom at plane $f(x, y, z = 128)$

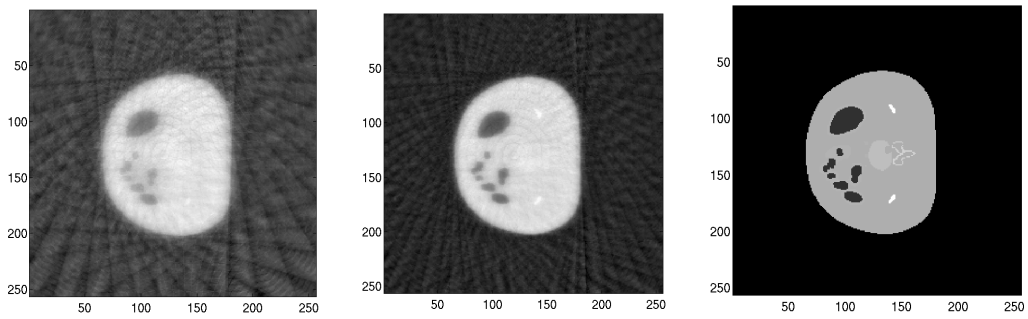
Figure 7.9: Cone beam (3D) reconstruction of the chest phantom.

leads to a better image reconstruction on a commodity computer.

7.3.2 Cone beam (3D) reconstruction of the chest phantom

In this section, the full 3D chest phantom was used for the reconstruction. Two cross-sections of the original image are shown in Fig. 7.9(c) and 7.10(c).

Fig. 7.9(a) shows the reconstruction by using LSQR to solve the LS problem obtained from only 18 scanned directions. Fig. 7.9(b) is the reconstruction by using LSQR with blockwise matrix-vector multiplications with 2 blocks of A (i.e. using 36 scanned directions). It is clear that the reconstructed image quality in Fig. 7.9(b) is better than the one in Fig. 7.9(a). More details were picked up in Fig. 7.9(b), while some of the details may even not appear in the reconstruction in Fig. 7.9(a).



(a) 3D reconstruction using 1/20 data (b) 3D reconstruction using 1/10 data (c) Chest phantom at plane $f(x, y, z = 60)$

Figure 7.10: Cone beam (3D) reconstruction of the chest phantom.

Fig. 7.10 is another set of reconstructed images for a different z plane of the chest phantom, where $f(x, y, z = 60)$. Once again, using more of the projection data leads to a better quality reconstruction.

Fig. 7.11 shows the relative residual norm plot for the reconstruction of Fig. 7.9(a) and 7.9(b). There is no significant difference in the two convergence curves. However, using more blocks of A ensures that LSQR converges to a more physically meaningful solution, as illustrated in Fig. 7.9. The plot in Fig. 7.11 also shows that in both cases the measured data is very noisy, as neither relative residual norm decreases beyond roughly 10^{-2} (even if many more than 10 LSQR iterations are performed). This has two important consequences. First, regardless of the algorithm used to solve the LS problem, it is not expected to generate an extremely accurate reconstruction. In addi-

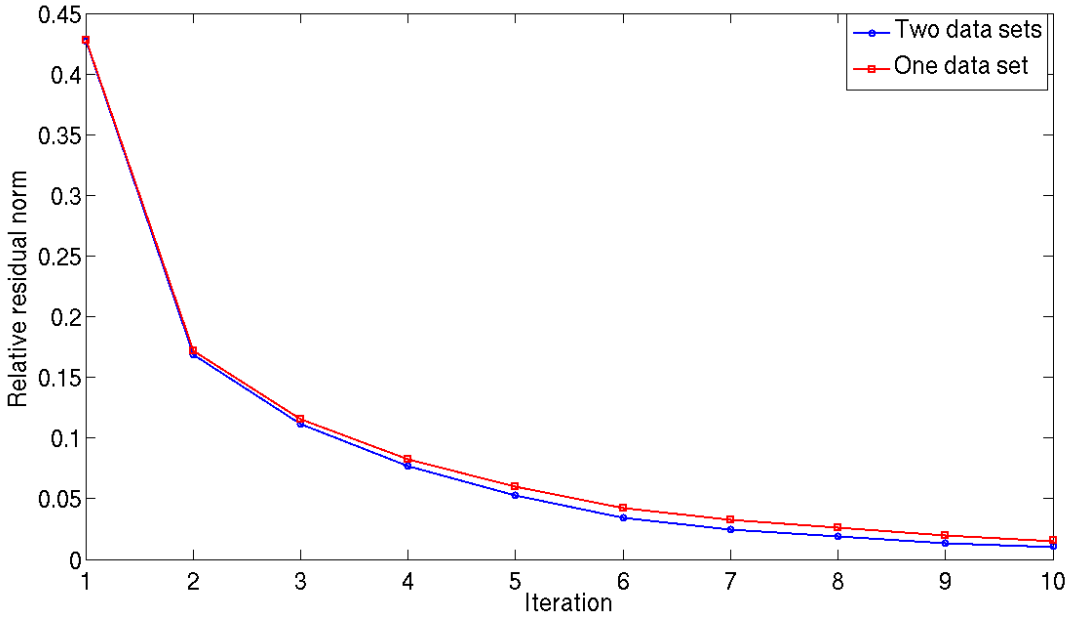


Figure 7.11: Relative residual norm plot for the reconstruction of Fig. 7.9.

tion, only very few iterations of LSQR are required to obtain a computed solution with as high accuracy as can be expected. This applies to the other test problems as well.

7.3.3 Cone beam (3D) reconstruction of the Rando head phantom

In this section, results are presented for the popular 'Rando' head phantom. For reference, the reconstructed image using the full data set and using the FDK-filtered back projection [8] algorithm is shown in Fig. 7.12. This reference image was reconstructed using the COBRA cone beam software developers' package from EXXIM (Exxim

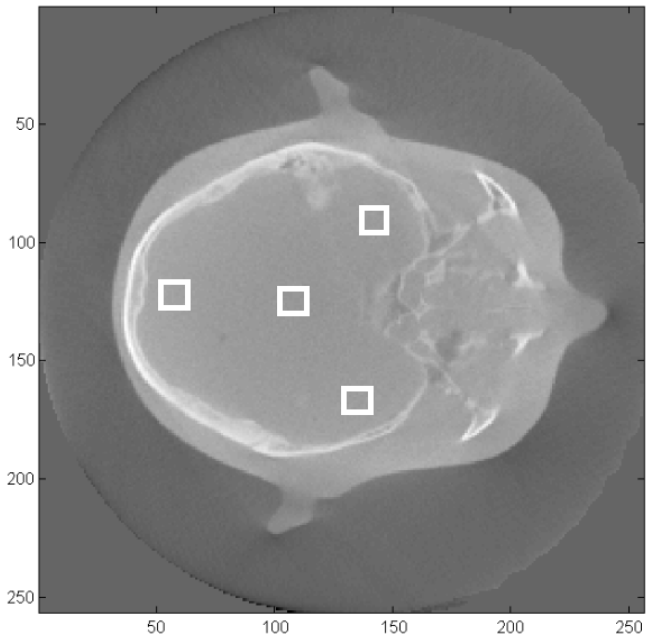
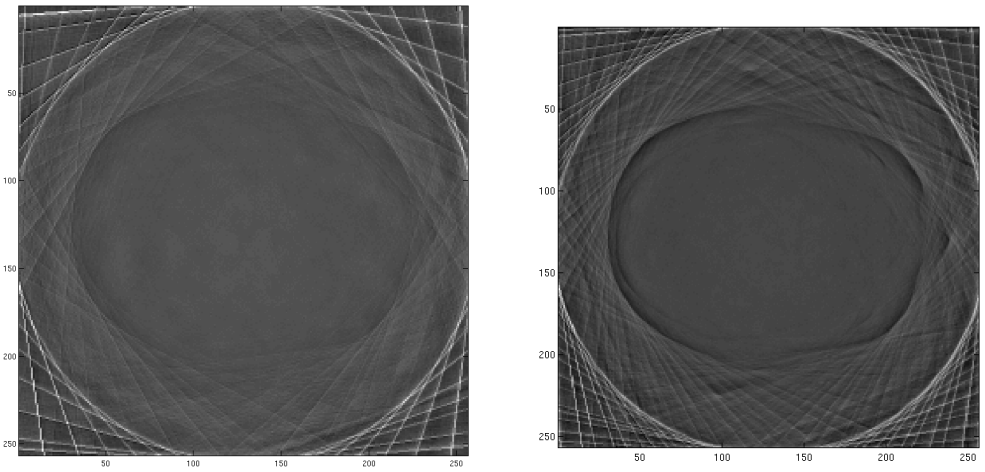


Figure 7.12: The experimental head image reconstructed with FDK FBP.

Computing Corporation, Pleasanton, CA, USA). Note that in the present work no additional filtering was applied to the reconstructed images. Therefore, it is not fair to expect the same quality image as the reference image.

Fig. 7.13 shows the reconstructions using LSQR with no regularization. When 1/20 of the available data is used for the reconstruction, as shown in Fig. 7.13(a), some of the head structure details do not even appear clearly. The reconstruction in Fig. 7.13(b) at least recovers the basic head structure.

Fig. 7.14 shows the relative residual norm plot in LSQR for the reconstruction in



(a) Experimental head reconstruction with LSQR with 1/20 of data

(b) Experimental head reconstruction with block LSQR with 1/10 of data

Figure 7.13: Examples of reconstructed images for Fig. 7.12.

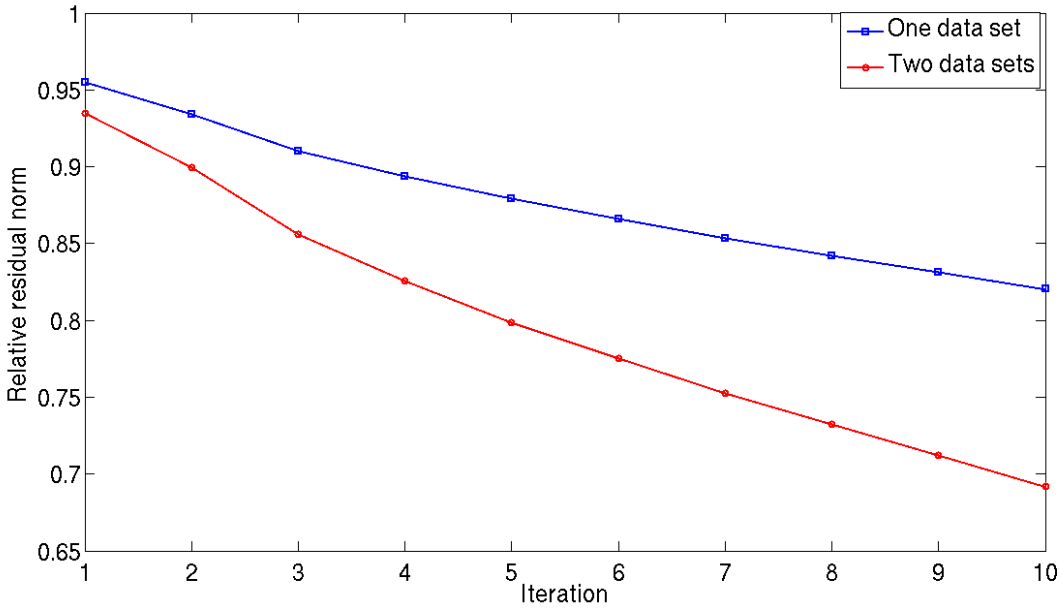


Figure 7.14: Relative residual norm plot for the reconstruction of Fig. 7.13.

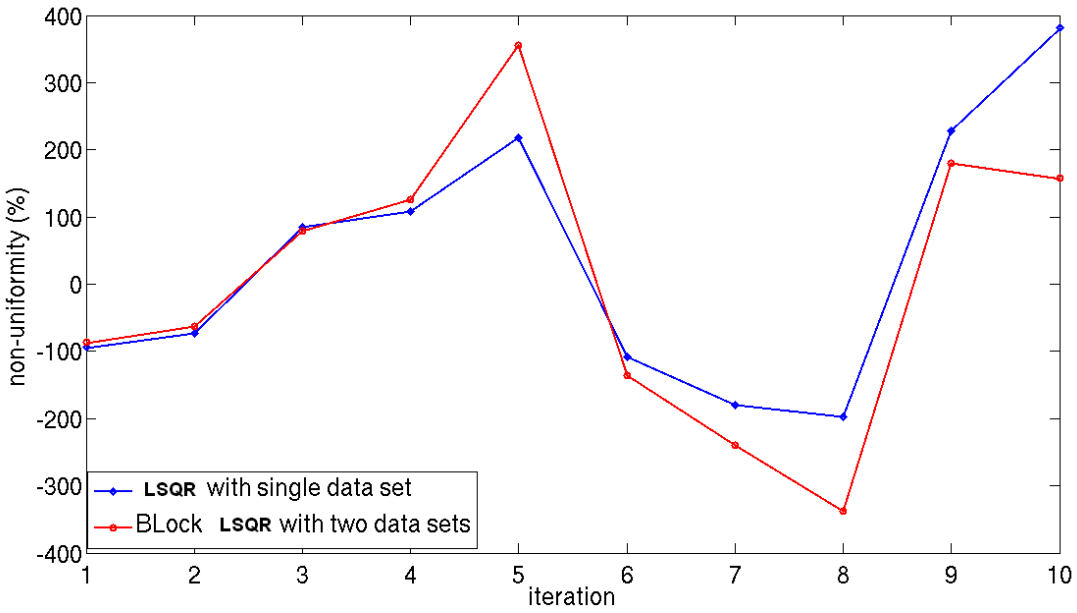


Figure 7.15: Comparison of uniformity.

Fig. 7.13(a) and 7.13(b). Again the rates of convergence are similar, but using more blocks of A ensures that the solution of the resulting LS problem to which LSQR converges is more physically meaningful.

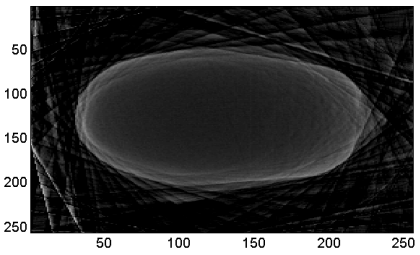
For the purpose of evaluation, image quality measurement (uniformity) is also introduced. As shown in Fig. 7.12, four regions of interest (ROI) are defined within the skull, which are expect to consist of a uniform density material. Hence, each of these regions is expected to have the same mean pixel value for a good quality image. The mean value for each ROI was calculated and uniformity was defined as the maximum

difference between any two ROIs quoted as a percentage of the central region mean value. Note that the uniformity value will be affected by scatter artefacts. This image quality measurement is plotted in Fig. 7.15. The final reconstructed image after 10 iterations has a lower non-uniformity when more blocks of A are used.

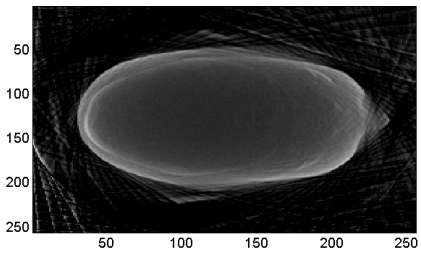
7.3.4 Regularization

Although implementing blockwise matrix-vector multiplications in LSQR provides a method to overcome memory capacity constraints and use more projection data for the reconstruction, the quality of reconstructed images may still not be very good in some cases. Fig. 7.16 shows reconstructions using LSQR with Tikhonov regularization. Compared with the reconstructions without regularization in Fig. 7.13, these reconstruction show significant improvement. Even when limited data are used, for instance in Fig. 7.16(b) and Fig. 7.16(c), many features found in the reference image are present in the reconstruction. Similar improvement have been observed using simulated data of the NCAT chest phantom. This confirms that regularized LSQR can be a better method for image reconstruction, in particular when limited data are available.

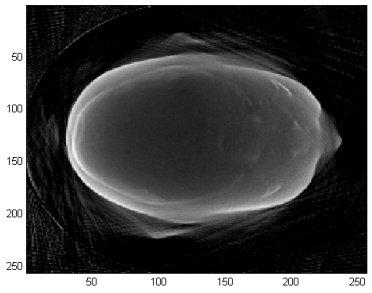
To investigate this further we have studied the reconstruction error and projection



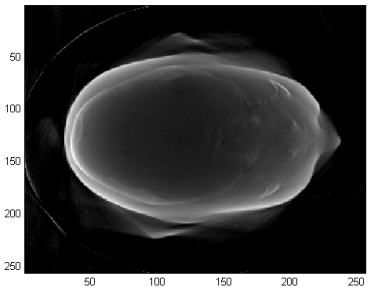
(a) 1/20 data



(b) 1/10 data



(c) 1/4 data



(d) Full data

Figure 7.16: Reconstruction of 1/20, 1/10, 1/4 and full data set in Rando in 3D.

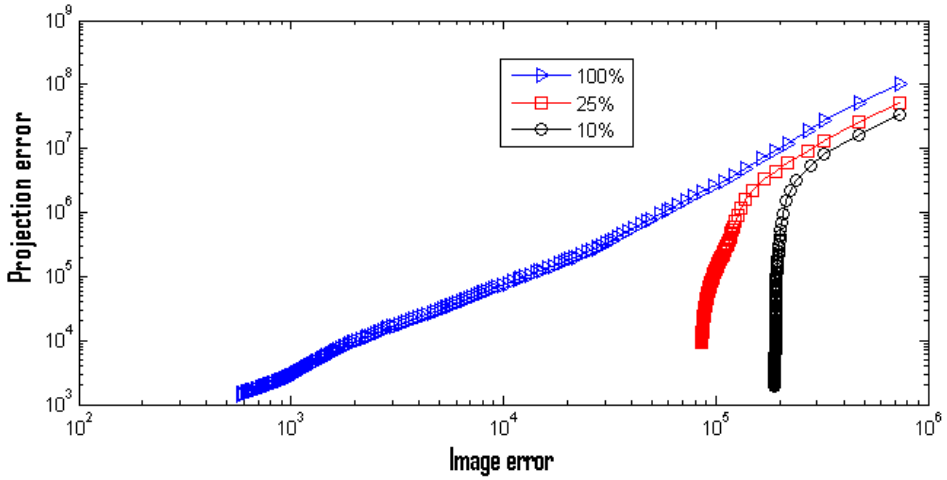


Figure 7.17: Image error vs projection error.

errors using simulated data from the NCAT phantom. Fig. 7.17 shows a plot of image error vs projection error for the iterative reconstruction of the NCAT phantom without the regularization. The image error is the norm of the difference between the true image in Fig. 7.8(d) and reconstructed image. Each data point in the plot corresponds to successive iterations of LSQR, from right to left, for a total of 100 iterations. The plot compares reconstructions using 100 percent, 25 percent, and 10 percent of the data. Although the projection error decreases in each case, the image error does not decrease as much when very limited data is used.

7.3.5 Computational Cost

Here some indication of computational time for the 3D image reconstruction is presented. The computer used is a 64-bit 3.33GHz microprocessor with 32GB of RAM running on a Linux operating system. The computational time for calculation of each projection matrix is 640 seconds, which can be done in advance and off-line. The Linux computer used had only 2 cores, and the matrix-vector multiplication in LSQR was carried out in parallel between the two cores. When 1/4 of all projection directions were used, the computational time for each iteration of LSQR was 1620 seconds, and 1720 seconds for regularized LSQR. It is worth mentioning that the computational time depends on number of projection directions used and also number of cores available to perform the matrix-vector products in parallel.

7.3.6 MIDAS, OS-SART and CGLS

In this section, the three main methods that are used in this thesis are compared with each other to show their differences while applying them to the same set of data, as shown in Fig. 7.18. It can be observed that, the convergence for MIDAS based iter-

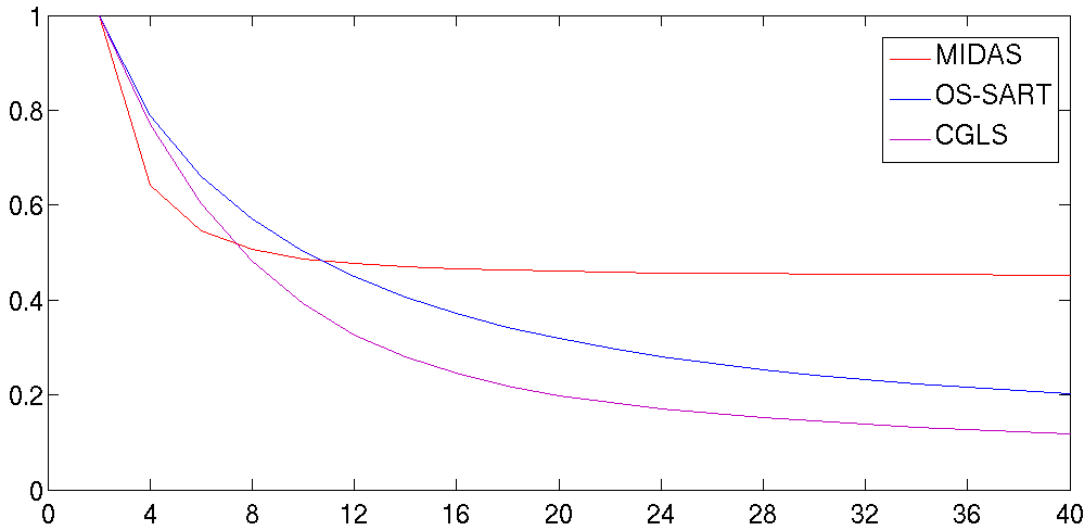
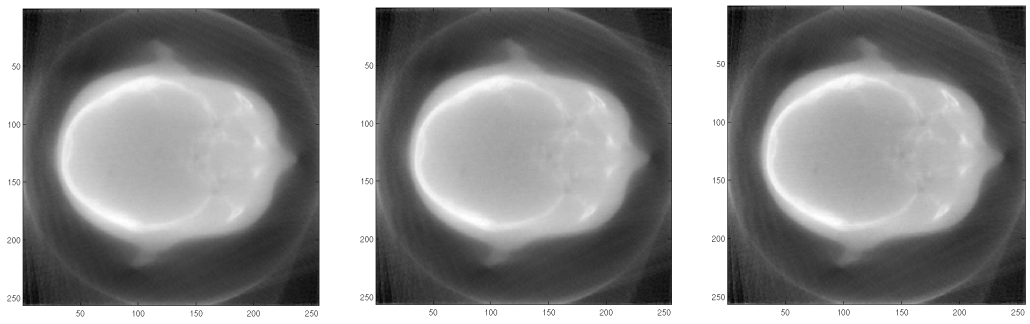


Figure 7.18: Convergence plot for the three methods applying to the same data set in 40 iterations.

ative method is much slower than the other two methods. The behaviours of the rest two methods are similar but CGLS converge a bit faster than OS-SART. In addition, visualised reconstructed images are shown in Fig.7.19. Here, CGLS presents a much better reconstruction.

7.4 Conclusion

The use of iterative methods in CBCT reconstruction has been well studied in the literature. The main challenge is that the amount of memory required to store the matrix A



(a) MIDAS (b) OS-SART (c) CGLS

Figure 7.19: Visualised image reconstructions.

is too large by commodity computer standards. Hence, the application of LSQR with blockwise matrix-vector multiplications in CBCT reconstruction was studied. The weighting matrix A has been partitioned into blocks, allowing each block to be loaded into memory sequentially or in parallel when needed at every iteration. This makes it possible to use the full matrix A for CBCT reconstruction. Partitioning the matrix A into blocks and performing the matrix-vector multiplications blockwise in LSQR allows us to use several blocks of A and gives a better reconstruction than using LSQR directly on a single block of A. The proposed blockwise implementation of LSQR allows the CBCT reconstruction using high performance computing and in particular parallel compactions. Partitioning the weighting matrix A according to computer memory capacity enables the use of CBCT reconstruction on commodity computers

with limited memory.

Furthermore, introducing regularization in LSQR can significantly improve the quality of imaging results. In this study standard Tikhonov regularization was used and produced satisfactory results. More advanced regularization, such as the method of total variation, will be investigated in our future studies. These may capture even more image features when limited data is used in the reconstruction.

Chapter 8

Conclusions

Cone beam computed tomography (CBCT) is widely used in many real life applications, as it provides volumetric and high quality image reconstruction from large scale projection data and it is especially useful in image guided radiation therapy (IGRT). In current commercial CBCT reconstruction systems, filtered back projection (FBP) like algorithm mainly based on the Feldkamp, Davis and Kress (FDK) algorithm [8] is routinely available in commercial radiology and radiotherapy equipment. In addition, another major reconstruction approach, iterative algorithms are also being investigated for clinical applications [14, 13]. Specifically, in this thesis, iterative algorithms for

CBCT reconstruction have been investigated. In this field, iterative algorithms have been well studied. However, in the research of CBCT image reconstruction, there is a main problem that even it is possible to access clinical projection data, the lack of a generalised iterative reconstruction package is an obstacle to researchers seeking to optimise or improve reconstructed image characteristics. MIDAS is a software package of this kind which is now being extended for iterative CBCT reconstruction in this project. Based on this platform, a lot of experiments and studies have been developed and investigated.

For iterative algorithms, though they were first proposed as an alternative approach for full-data tomographic image reconstruction, studies were found that they were also applicable, and provide better performance for down-sampled data or limited angle tomographic reconstruction [27, 26, 18]. And recently, work has shown that prior image information can assist reconstructions where significant amounts of projection data are missing [34, 13]. In iterative algorithms, algorithm parameters are important factors affecting the performance of the reconstruction. However, the question about how to evaluate the clinical image using iterative algorithms is still an open question due to the special requirements of medical applications, although some criteria have

been discussed in [36, 2, 37]. Depending on the particular application, there is always a trade-off between different reconstruction algorithm parameters, especially for the relaxation factor of the algorithms.

In IGRT applications, it is desirable to lower the patient radiation dose while maintaining good quality tomographic reconstructions. Hence, methods are needed even when the data is undersampled or there are limited projections. Iterative algorithms such as ART, SART and OS-SART are known to perform well under such circumstances. The performance of ART, SART and OS-SART is studied in this projection based on a range of norm measurements (RMS reconstruction error, RMS projection error, and the L1 norm and L2 norm of the reconstruction). Image quality measurements are also introduced. Since image quality often degrades as iterative algorithms converge, a function is introduced and used to trade off convergence with image quality.

In iterative algorithms, LSQR [59, 60] is a well-known iterative method for solving (possibly regularized) large sparse non-symmetric systems of equations $Ax = b$ and linear least squares problems $\min_x \|b - Ax\|_2$. Although LSQR (or other implementations of CG on the normal equations) is often considered the method of choice

for the iterative solution of linear least squares problems, the use of such methods in tomography problems is not widespread. Hence, CBCT reconstruction using LSQR is studied in this project. The main challenge in CBCT reconstruction is that the weighting matrix A is usually very large that storing it in full requires an amount of memory well beyond the reach of typical commodity computer. This also applies when using CG-type methods for CBCT reconstruction. In the thesis, attempts have been made to overcome this shortcoming. All that required of the matrix A in CG-type methods such as LSQR is a subroutine to perform matrix-vector multiplications by A and $A^\top$. The matrix A can be partitioned in a blockwise manner, and each block can be loaded into memory sequentially or in parallel (depending on the computer platform available) during the matrix-vector multiplications. This implementation makes it possible to use the full weighting matrix A for CBCT reconstruction, without further enhancing computer standards.

The main contributions of this thesis are as follows:

- A general Matlab platform is developed, specifically for the research of iterative algorithms for CBCT image reconstruction. Various algorithms are im-

plemented, including ART, SART, OS-SART and conjugate gradient type algorithm.

- Parameters optimisation are investigated for ART algorithm. The settings of relaxation parameter and number of iterations are optimised based on reconstruction stability, convergence rate and image quality. For different measurement criteria, the 1D image row plot shows the quality of the reconstructed image and contrast of edges, while the transverse sum plot focuses on reconstruction for a specific area. With these evaluations, even for down sampled projection data such as 1/2, 2/3, 3/4 and 4/5, better reconstructions can be achieved by applying the optimisation approach.

- A practical study was conducted based on theoretical results in convergence for different iterative algorithms (ART, SART and OS-SART), focusing on convergence properties and image quality parameters (uniformity and noise) and OS-SART is investigated more thoroughly. With the data both for 2D (simulated and experimental) and 3D (CBCT), it suggests that from a convergence point of view, for OS-SART, the larger the block size, the faster the convergence. However im-

age quality typically degrades as the algebraic algorithms converge. A trade-off between convergence and image quality should be made to maximise the performance of iterative algorithms. Hence, a function is introduced to implement this trade-off.

- In order to overcome the challenge of memory requirement to store the matrix A in CBCT image reconstruction. The application of LSQR with blockwise matrix-vector multiplications is studied. The weighting matrix A has been partitioned into blocks, allowing each block to be loaded into memory sequentially or in parallel when needed at every iteration, enabling the use of full matrix A for reconstruction. In addition, the proposed approach allows the implementation of high performance computing and parallel compactions. Furthermore, by using this approach, commodity computers with limited memory can also implement CBCT reconstruction.

- Regularised version of LSQR is investigated. It can significantly improve the quality of imaging results, especially for the case of image reconstruction of down sampled data.

Future work

Iterative algorithms have been employed in many areas as they provide the potential of generating high quality image reconstruction and they are particularly useful for tomographic imaging in the case that when projection data are insufficient. However, there are still a lot can be done to improve these methods for medical use purpose.

- High performance computing like GPU are proposed to overcome the huge amount of time cost during forward and backward projection during iterations. With GPU acceleration, the speed of computational time cost can reduce by a factor of 40-100 times.

- Incomplete data of CBCT occur frequently in medical imaging. The incomplete data problems in which projection data are only available in an angular range can be attributed to the limited angle tomography. Iterative algorithms have been demonstrated to perform better than traditional FBP technique for limited angle tomography.

- When applied to small animal imaging and other similar problems, Micro-CT is used. Iterative algorithms have been studied in CBCT for many years and may

also apply successfully in micro-CT. However, time cost of iterative algorithm is still the main obstacle for its use even for micro-CT.

- Compressed sensing allows to reduce a large amount of data while it also enables to recover the data following its rules. Combining iterative algorithms and compressed sensing will save a hugh amount of time during the forward and backward projection in iterative reconstruction process.

Publications

Journal papers:

1. **Qiu, W.**, Tong, J., Mitchell, C., Marchant, T., Spencer, P., Moore, C.J. and Soleimani, M., 2010. New iterative cone beam CT reconstruction software: parameter optimisation and convergence study. *Computer Methods and Programs in Biomedicine*, 100(2), pp.166-174. (ISI journal impact factor 1.24).

2. Pengpan, T., Smith, N.D., **Qiu, W.**, Yao, A., Mitchell, C.N. and Soleimani, M., 2011. A motion-compensated cone-beam CT using electrical impedance tomography imaging. *Physiological Measurement*, 32(1), pp.19-34. (ISI journal impact factor 1.57).

3. Pengpan, T., **Qiu, W.**, Smith, N.D. and Soleimani, M. Cone-beam CT using motion-compensated algebraic reconstruction methods with limited data, *In press for Computer Methods and Programs in Biomedicine*, (ISI journal impact factor 1.24).

4. **Qiu, W.**, Titley-Peloquin, D. and Soleimani M, Blockwise conjugate gradient methods for image reconstruction in volumetric CT, *In press for Computer Methods and Programs in Biomedicine*, (ISI journal impact factor 1.24).

5. **Qiu, W.**, Pengpan, T., Smith, N.D. and Soleimani, M., Evaluating iterative algebraic algorithms in terms of convergence and image quality for cone beam CT, *Submitted for Computer Methods and Programs in Biomedicine*, (ISI journal impact factor 1.24).

Conference papers:

1. **Qiu, W.**, W., Soleimani, M., Mitchell, C.N., Marchant, T. and Moore, C.J., 2010. Iterative image reconstruction methods in Cone Beam CT applied to phantom and clinical data. *In: International Conference on Computer Vision: Theory and Applications (VISAPP 2010)*, 17-21 May 2010, Angers, France.

2. Soleimani, M., **Qiu, W.**, W. and Mitchell, C., 2010. A Modular tomography software for soft field and hard field tomography. *In: 6th World Congress on Industrial Process Tomography (WCIPT6)*, 6-9 September 2010, Beijing, China.

References

[1] K. Mueller. *Fast and accurate three-dimensional reconstruction from cone-beam projection data using algebraic methods*. PhD thesis, The Ohio State University, 1998.

[2] A. C. Kak and M. Slaney. *Principles of computerized tomographic imaging*. IEEE Press, New York, 1988.

[3] J. Radon. On the Determination of Functions from Their Integral Values along Certain Manifolds. *IEEE Transactions on Medical Imaging*, 5(4):170–176, Dec 1986.

[4] R. Bracewell and P. B. Kahn. The Fourier Transform and Its Applications. *American Journal of Physics*, 34(8):712–712, 1966.

[5] R. N. Bracewell and A. C. Riddle. Inversion of Fan-Beam Scans in Radio Astronomy. *Astrophysics Journal*, 150:427–434, Nov 1967.

[6] G. N. Ramachandran and A. V. Lakshminarayanan. Three-dimensional Reconstruction from Radiographs and Electron Micrographs: Application of Convolutions instead of Fourier Transforms. *Proceedings of the National Academy of Sciences of the United States of America*, 68(9):2236–2240, 1971.

[7] L. A. Shepp and B. F. Logan. The Fourier reconstruction of a head section. *IEEE Transactions on Unclear Science*, NS21(3):21–43, 1974.

[8] L. A. Feldkamp, L. C. Davis, and J. W. Kress. Practical cone-beam algorithm. *Journal of the Optical Society of America A*, 1(6):612–619, 1984.

[9] R. Gordon, R. Bender, and G. T. Herman. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography. *Journal of Theoretical Biology*, 29(3):471–481, 1970.

[10] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39(1):1–38, 1977.

[11] P. Gilbert. Iterative methods for the three-dimensional reconstruction of an object from projections. *Journal of Theoretical Biology*, 36(1):105–117, 1972.

[12] A. H. Andersen and A. C. Kak. Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm. *Ultrasonic imaging*, 6(1):81–94, 1984.

[13] G. Chen, J. Tang, and S. Leng. Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Medical Physics*, 35(2):660–663, 2008.

[14] J. Wang, T. Li, and L. Xing. Iterative image reconstruction for CBCT using edge-preserving prior. *Medical Physics*, 36(1):252–260, 2009.

[15] X. Pan, E. Y. Sidky, and M. Vannier. Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction? *Inverse Problems*, 25(12):123009, 2009.

[16] M. Trummer. A note on the ART of relaxation. *Computing*, 33(3):349–352, Sep 1984.

[17] M. Jiang and G. Wang. Convergence studies on iterative algorithms for image reconstruction. *IEEE Transactions on Medical Imaging*, 22(5):569–579, May 2003.

[18] K. Mueller, R. Yagel, and J. J. Wheller. Anti-aliased three-dimensional cone-beam reconstruction of low-contrast objects with algebraic methods. *IEEE Transactions on Medical Imaging*, 18(6):519–537, 1999.

[19] C. N. Mitchell and P. Spencer. A three-dimensional time-dependent algorithm for ionospheric imaging using GPS. *Annals of Geophysics*, 46(4):687–696, Aug 2003.

[20] T. E. Marchant, C. J. Moore, C. G. Rowbottom, R. I. MacKay, and P. C. Williams. Shading correction algorithm for improvement of cone-beam CT images in radiotherapy. *Physics in Medicine and Biology*, 53(20):5719–5733, 2008.

[21] C. J. Moore, A. Amer, T. Marchant, J. R. Sykes, J. Davies, J. Stratford, C. McCarthy, C. MacBain, A. Henry, P. Price, and P. C. Williams. Developments in and experience of kilovoltage X-ray cone beam image-guided radiotherapy. *British Journal of Radiology*, 79(Special Issue 1):S66–78, 2006.

[22] S. Haykin. *Array signal processing*. Prentice Hall, New Jersey, 1985.

[23] M. E. Davison and F. A. Grunbaum. Convolution algorithms for arbitrary projection angles. *IEEE Transactions on Nuclear Science*, 26(2):2670–2673, Apr 1979.

[24] M. Nassi, W. R. Brody, B. P. Medoff, and A. Macovski. Iterative reconstruction reprojection: an algorithm for limited data cardiac-computed tomography. *IEEE Transactions on Biomedical Engineering*, BME-29(5):333–341, May 1982.

[25] J. A. Reeds and L. A. Shepp. Limited angle reconstruction in tomography via squashing. *IEEE Transactions on Medical Imaging*, 6(2):89–97, Jun 1987.

[26] A. H. Andersen. Algebraic reconstruction in CT from limited views. *IEEE Transactions on Medical Imaging*, 8(1):50–55, Mar 1989.

[27] R. Rangayyan, A. P. Dhawan, and R. Gordon. Algorithms for limited-view computed tomography: an annotated bibliography and a challenge. *Applied Optics*, 24(23):4000–4012, 1985.

[28] B. P. Medoff, W. R. Brody, M. Nassi, and A. Macovski. Iterative convolution backprojection algorithms for image reconstruction from limited data. *Journal of the Optical Society of America*, 73(11):1493–1500, 1983.

[29] J. H. Kim, K. Y. Kwak, S. B. Park, and Z. H. Cho. Projection space iteration reconstruction-reprojection. *IEEE Transactions on Medical Imaging*, 4(3):139–143, Sep 1985.

[30] K. Lange and J. A. Fessler. Globally convergent algorithms for maximum a posteriori transmission tomography. *IEEE Transactions on Image Processing*, 4(10):1430–1438, Oct 1995.

[31] L. Parra and H. H. Barrett. List-mode likelihood: EM algorithm and image quality estimation demonstrated on 2-D PET. *IEEE Transactions on Medical Imaging*, 17(2):228–235, April 1998.

[32] J.-B. Thibault, K. Sauer, C. Bouman, and Hsieh J. High quality iterative image reconstruction for multi-slice helical CT. In *Proceedings of the 3D'03 Conference*, 2003.

[33] Y. Z. O'Connor and J. A. Fessler. Fourier-based forward and back-projectors in iterative fan-beam tomographic image reconstruction. *IEEE Transactions on Medical Imaging*, 25(5):582–589, 2006.

[34] E. J. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, Feb 2006.

[35] M. Jiang and G. Wang. Development of iterative algorithms for image reconstruction. *Journal of X-Ray Science and Technology*, 10(1):77–86, Jan 2001.

[36] G. T. Herman. *Image reconstruction from projections : the fundamentals of computerized tomography*. Computer science and applied mathematics. Academic Press, San Francisco, 1980.

[37] J. L. Prince and J. M. Links. *Medical imaging signals and systems*. Prentice Hall, New Jersey, 2006.

[38] G. T. Herman and L. B. Meyer. Algebraic reconstruction techniques can be made computationally efficient. *IEEE Transactions on Medical Imaging*, 12(3):600–609, sep 1993.

[39] M. Soleimani, W. Qiu, and C.N. Mitchell. A modular tomography software for soft field and hard field tomography. In *6th World Congress on Industrial Process Tomography (WCIPT6), Beijing, China*, 2010.

[40] T. Strohmer and R. Vershynin. A Randomized Kaczmarz Algorithm with Exponential Convergence. *Journal of Fourier Analysis and Applications*, 15:262–278, 2009.

[41] D. Needell. Randomized Kaczmarz solver for noisy linear systems. *BIT Numerical Mathematics*, 50:395–403, 2010.

[42] W. Qiu, J. R. Tong, C. N. Mitchell, T. Marchant, P. Spencer, C. J. Moore, and M. Soleimani. New iterative cone beam CT reconstruction software: Parameter optimisation and convergence study. *Computer Methods and Programs in Biomedicine*, 100(2):166–174, Nov 2010.

[43] M. R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *Journal cf Research of the National Bureau of Standards*, 49(6):409 – 436, 1952.

[44] A. van der Sluis and H. A. van der Vorst. The rate of convergence of Conjugate Gradients. *Numerische Mathematik*, 48:543–560, 1986. 10.1007/BF01389450.

[45] J. R. Shewchuk. An introduction to the conjugate gradient method without the agonizing pain. Technical report, Carnegie Mello University, Pittsburgh, PA, USA, 1994.

[46] P. C. Hansen. *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*. Society for Industrial and Applied Mathematics, 1998.

[47] Y. Saad. *Iterative methods for sparse linear systems*. Society for Industrial and Applied Mathematics, 2003.

[48] S. Kawata and O. Nalcioglu. Constrained iterative reconstruction by the conjugate gradient method. *IEEE Transactions on Medical Imaging*, 4(2):65 – 71, 1985.

[49] J. A. Scales. Tomographic inversion via the conjugate gradient method. *Geophysics*, 52(2):179–185, 1987.

[50] J. M. Lees and R. S. Crosson. Bayesian Art versus Conjugate Gradientf Methods in Tomographic Seismic Imaging: An Application at Mount St. Helens, Washington. *Lecture Notes-Monograph Series*, 20:pp. 186–208, 1991.

[51] P. Calvini and M. Bertero. Application of Krylov subspaces to SPECT imaging. *International Journal of Imaging Systems and Technology*, 12(5):217 – 228, 2002.

[52] E. L. Piccolomini and F. Zama. Conjugate gradient regularization method in Computed Tomography problems. *Applied Mathematics and Computation*, 102(1):87 – 99, 1999.

[53] M. Wang. Inverse solutions for electrical impedance tomography based on conjugate gradients methods. *Measurement Science and Technology*, 13(1):101 – 117, 2002.

[54] Y. Zhang, C. Yu, and D. Chen. A novel image reconstruction method based on weighted SVD truncation conjugate gradient algorithm for electrical resistance tomography system. *Journal of Computational Information Systems*, 6(2):395 – 404, 2010.

[55] R. Fahrig and D. W. Holdsworth. Three-dimensional computed tomographic reconstruction using a C-arm mounted XRII: Image-based correction of gantry motion nonidealities. *Medical Physics*, 27(1):30–38, 2000.

[56] R. M. Lewitt. Alternatives to voxels for image representation in iterative reconstruction algorithms. *Physics in Medicine and Biology*, 37(3):705, 1992.

[57] A. Ziegler, T. Kohler, T. Nielsen, and R. Proksa. Efficient projection and backprojection scheme for spherically symmetric basis functions in divergent beam geometry. *Medical Physics*, 33(12):4653–4663, 2006.

[58] T. Pengpan, N. D. Smith, W. Qiu, A. Yao, C. N. Mitchell, and M. Soleimani. A motion-compensated cone-beam CT using electrical impedance tomography imaging. *Physiological Measurement*, 32(1):19–34, Jan 2011.

[59] C. C. Paige and M. A. Saunders. LSQR: an algorithm for sparse linear equations and sparse least squares. *ACM Transactions on Mathematical Software*, 8(1):43–71, 1982.

[60] C. C. Paige and M. A. Saunders. Algorithm 583, LSQR:sparse linear equations and sparse least squares problems. *ACM Transactions on Mathematical Software*, 8(2):195–209, 1982.

[61] P. Koken and M. Grass. Aperture weighted cardiac reconstruction for cone-beam CT. *Physics in Medicine and Biology*, 51(14):3433, 2006.

[62] A. Ziegler, T. Nielsen, and M. Grass. Iterative reconstruction of a region of interest for transmission tomography. *Medical Physics*, 35(4):1317–1327, 2008.

[63] A. A. Isola, A. Ziegler, T. Koehler, W. J. Niessen, and M. Grass. Motion-compensated iterative cone-beam CT image reconstruction with adapted blobs as basis functions. *Physics in Medicine and Biology*, 53(23):6777, 2008.

[64] G. H. Golub and C. F. Van Loan. *Matrix Computations (Johns Hopkins Studies in Mathematical Sciences)(3rd Edition)*. The Johns Hopkins University Press, 3rd edition, October 1996.

[65] W. P. Segars. *Development and Application of the New Dynamic NURBS-based Cardiac-Torso (NCAT) Phantom*. PhD thesis, University of North Carolina, 2001.

[66] M. Bertram, J. Wiegert, D. Schafer, T. Aach, and G. Rose. Directional View Interpolation for Compensation of Sparse Angular Sampling in Cone-Beam CT. *IEEE Transactions on Medical Imaging*, 28(7):1011–1022, July 2009.

[67] C. Byrne. A unified treatment of some iterative algorithms in signal processing and image reconstruction. *Inverse Problems*, 20(1):103, 2004.

[68] A. Dax. An Open Question on Cyclic Relaxation. *BIT Numerical Mathematics*, 43:929–943, 2003.

[69] J. T. Dobbins III and D. J. Godfrey. Digital x-ray tomosynthesis: current state of the art and clinical potential. *Physics in Medicine and Biology*, 48(19):65–106, 2003.

[70] J. Gregor and T. Benson. Computational analysis and improvement of SIRT. *IEEE Transactions on Medical Imaging*, 27(7):918–924, 2008.

[71] G. T. Herman and S. S. Trivedi. Comparative study of two postreconstruction beam hardening correction methods. *IEEE Transactions on Medical Imaging*, MI-2(3):128–135, 1983.

[72] M. Kachelriess, K. Sourbelle, and W. A. Kalender. Empirical cupping correction: a first-order raw data precorrection for cone-beam computed tomography. *Medical Physics*, 33(5):1269–1274, 2006.

[73] S. Kaczmarz. Angenherte Auflsung von Systemen linearer Gleichungen. *Bulletin International de l'Academie Polonaise des Sciences et des Lettres*, 35:335–357, 1937.

[74] J. Kaipio and E. Somersalo. *Statistical and computational inverse problems*. Springer, New York, 2005.

[75] A. C. Kak. Computerized tomography with X-ray, emission, and ultrasound sources. *Proceedings of the IEEE*, 67(9):1245–1272, 1979.

[76] T. E. Marchant, A. M. Amer, and C. J. Moore. Measurement of inter and intra fraction organ motion in radiotherapy using cone beam CT projection images. *Physics in Medicine and Biology*, 53(4):1087–1098, 2008.

[77] C. J. Moore, T. E. Marchant, and A. M. Amer. Cone beam CT with zonal filters for simultaneous dose reduction, improved target contrast and automated set-up in radiotherapy. *Physics in Medicine and Biology*, 51(9):2191–2204, 2006.

[78] K. Mueller, R. Yagel, and J. J. Wheller. Fast implementations of algebraic methods for three-dimensional reconstruction from cone-beam data. *IEEE Transactions on Medical Imaging*, 18(6):538–548, 1999.

[79] F. Natterer. *The mathematics of computerized tomography*. SIAM, 2001.

[80] T. Pengpan, W. Qiu, N. D. Smith, and M. Soleimani. Cone-beam CT using motion-compensated algebraic reconstruction methods with limited data. *Computer Methods and Programs in Biomedicine (in press)*.

[81] T. Pengpan, N. D. Smith, W. Qiu, A. Yao, C. N. Mitchell, and M. Soleimani. A motion-compensated cone-beam CT using electrical impedance tomography imaging. *Physiological Measurement*, 32 (1):19–34, 2011.

[82] C. Popa and R. Zdunek. Kaczmarz extended algorithm for tomographic image reconstruction from limited-data. *Mathematics and Computers in Simulation*, 65(6):579 – 598, 2004.

[83] W. Qiu, T. Pengpan, N. D. Smith, and M. Soleimani. Evaluating iterative algebraic algorithms in terms of convergence and image quality for cone beam CT. *International Journal of Numerical Analysis & Modeling - Series B (submitted)*.

[84] W. Qiu and M. Soleimani. A modular tomography software for soft field and hard field tomography. *Chinese Journal of Chemical Engineering (submitted)*.

[85] W. Qiu, M. Soleimani, C. N. Mitchell, T. Marchant, and C. J. Moore. Iterative image reconstruction methods in Cone Beam CT applied to phantom and clinical

data. In *International Conference on Computer Vision: Theory and Applications (VISAPP 2010), Angers, France*, 17-21 May 2010.

[86] W. Qiu, D. Titley-Peloquin, and M. Soleimani. Blockwise conjugate gradient methods for image reconstruction in volumetric CT. *Computer Methods and Programs in Biomedicine (submitted)*.

[87] W. Qiu, J. Tong, C. Mitchell, T. Marchant, P. Spencer, C. J. Moore, and M. Soleimani. New iterative cone beam CT reconstruction software: parameter optimisation and convergence study. *Computer Methods and Programs in Biomedicine*, 100 (2):166–174, 2010.

[88] W. Qiu, J. R. Tong, C. N. Mitchell, T. Marchant, P. Spencer, C. J. Moore, and M. Soleimani. Iterative image reconstruction methods in Cone Beam CT applied to phantom and clinical data. In *International Conference on Computer Vision: Theory and Applications (VISAPP 2010), Angers, France*, 2010.

[89] M. Rantala, S. Vanska, S. Jarvenpaa, M. Kalke, M. Lassas, J. Moberg, and S. Siltanen. Wavelet-based reconstruction for limited-angle X-ray tomography. *IEEE Transactions on Medical Imaging*, 25(2):210–217, Feb 2006.

[90] K. F. Riley. *Mathematical methods for physics and engineering*. Cambridge University Press, 2006.

[91] E. Y. Sidky, C. Kao, and X. Pan. Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT. *Journal of X-Ray Science and Technology*, 14(2):119–139, Jan 2006.

[92] M. Soleimani, W. Qiu, and C. Mitchell. A Modular tomography software for soft field and hard field tomography. In *6th World Congress on Industrial Process Tomography (WCIPT6), Beijing, China*, 6-9 September 2010.

[93] M. Soleimani, W. Qiu, and C. Mitchell. A modular tomography software for soft field and hard field tomography. In *6th World Congress on Industrial Process Tomography (WCIPT6), Beijing, China*, 2010.

[94] G Wang and M. Jiang. Ordered-subset simultaneous algebraic reconstruction techniques (OS-SART). *Journal of X-Ray Science and Technology*, 12(3):169–177, 2004.

[95] D. L. Wilson, A. J. Baddeley, and R. A. Owens. New metric for grey-scale image comparison. *International Journal of Computer Vision*, 24(1):5–17, 1997.

[96] F. Xu and K. Mueller. Accelerating popular tomographic reconstruction algorithms on commodity PC graphics hardware. *IEEE Transactions on Nuclear Science*, 52(3):654–663, June 2005.

[97] F. Xu and K. Mueller. Ultra-fast 3D filtered backprojection on commodity graphics hardware. In *IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2004*, pages 571–574 Vol. 1, April 2004.

[98] F. Xu and K. Mueller. Towards a unified framework for rapid 3D computed tomography on commodity GPUs. In *IEEE Nuclear Science Symposium Conference Record, 2003*, volume 4, pages 2757–2759 Vol.4, Oct. 2003.